Portfolio Optimization Problems:
A Martingale and a Convex Duality Approach

by

Nicole Flaure Kouemo Tchamga

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics at the University of Stellenbosch

Department of Mathematical Sciences
University of Stellenbosch
Private Bag X1, 7602 Matieland, South Africa

Supervisor: Dr. R. Ghomrasni

December 2010

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work,that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

------------------------------   ------------------------

Nicole Flaure Kouemo Tchamga   Date

Abstract

The first approach initiated by Merton [Mer69, Mer71] to solve utility maximization portfolio problems in continuous time is based on stochastic control theory. The idea of Merton was to interpret the maximization portfolio problem as a stochastic control problem where the trading strategies are considered as a control process and the portfolio wealth as the controlled process. Merton derived the Hamilton-Jacobi-Bellman (HJB) equation and for the special case of power, logarithm and exponential utility functions he produced a closed-form solution. A principal disadvantage of this approach is the requirement of the Markov property for the stocks prices. The so-called *martingale method* represents the second approach for solving utility maximization portfolio problems in continuous time. It was introduced by Pliska [Pli86], Cox and Huang [CH89, CH91] and Karatzas et al. [KLS87] in different variant. It is constructed upon convex duality arguments and allows one to transform the initial dynamic portfolio optimization problem into a static one and to resolve it without requiring any "Markov" assumption. A definitive answer (necessary and sufficient conditions) to the utility maximization portfolio problem for terminal wealth has been obtained by Kramkov and Schachermayer [KS99]. In this thesis, we study the convex duality approach to the expected utility maximization problem (from terminal wealth) in continuous time stochastic markets, which as already mentioned above can be traced back to the seminal work by Merton [Mer69, Mer71]. Before we detail the structure of our thesis, we would like to emphasize that the starting point of our work is based on Chapter 7 in Pham [P09] a recent textbook. However, as the careful reader will notice, we have deepened and added important notions and results (such as the study of the upper (lower) hedge, the characterization of the essential supremum of all the possible prices, compare Theorem 7.2.2 in Pham [P09] with our stated Theorem 2.4.9, the dynamic programming equation 2.31, the superhedging theorem 2.6.1...) and we have made a considerable effort in the proofs. Indeed, several proofs of theorems in Pham [P09] have serious gaps (not to mention typos) and even flaws (for example see the proof of Proposition 7.3.2 in Pham

[P09] and our proof of Proposition 3.4.8). In the first chapter, we state the expected utility maximization problem and motivate the convex dual approach following an illustrative example by Rogers [KR07, R03]. We also briefly review the von Neumann - Morgenstern Expected Utility Theory. In the second chapter, we begin by formulating the superreplication problem as introduced by El Karoui and Quenez [KQ95]. The fundamental result in the literature on super-hedging is the dual characterization of the set of all initial endowments leading to a super-hedge of a European contingent claim. El Karoui and Quenez [KQ95] first proved the superhedging theorem 2.6.1 in an Itô diffusion setting and Delbaen and Schachermayer [DS95, DS98] generalized it to, respectively, a locally bounded and unbounded semimartingale model, using a Hahn-Banach separation argument. The superreplication problem inspired a very nice result, called the optional decomposition theorem for supermartingales 2.4.1, in stochastic analysis theory. This important theorem introduced by El Karoui and Quenez [KQ95], and extended in full generality by Kramkov [Kra96] is stated in Section 2.4 and proved at the end of Section 2.7. The third chapter forms the theoretical core of this thesis and it contains the statement and detailed proof of the famous Kramkov-Schachermayer Theorem that addresses the duality of utility maximization portfolio problems. Firstly, we show in Lemma 3.2.1 how to transform the dynamic utility maximization problem into a static maximization problem. This is done thanks to the dual representation of the set of European contingent claims, which can be dominated (or super-hedged) almost surely from an initial endowment x and an admissible self-financing portfolio strategy given in Corollary 2.5 and obtained as a consequence of the optional decomposition of supermartingale. Secondly, under some assumptions on the utility function, the existence and uniqueness of the solution to the static problem is given in Theorem 3.2.3. Because the solution of the static problem is not easy to find, we will look at it in its dual form. We therefore synthesize the dual problem from the primal problem using convex conjugate functions. Before we state the Kramkov-Schachermayer Theorem 3.4.1, we present the Inada Condition and the Asymptotic Elasticity Condition for Utility functions. For the sake of clarity, we divide the long and technical proof of Kramkov-Schachermayer Theorem 3.4.1 into several lemmas and propositions of independent interest, where the required assumptions are clearly indicate for each step of the proof. The key argument in the proof of Kramkov-Schachermayer Theorem is an infinite-dimensional version of the minimax theorem (the classical method of finding a saddlepoint for the Lagrangian is not enough in our situation), which is central in the theory of Lagrange multipliers. For this, we have stated and proved the technical Lemmata 3.4.5 and 3.4.6. The main steps in the proof of the the Kramkov-Schachermayer Theorem 3.4.1 are:

- We show in Proposition 3.4.9 that the solution to the dual problem exists and we characterize it in Proposition 3.4.12.
- From the construction of the dual problem, we find a set of necessary and sufficient conditions (3.1.1), (3.1.2), (3.3.1) and (3.3.7) for the primal and dual problems to each have a solution.
- Using these conditions, we can show the existence of the solution to the given problem and characterize it in terms of the market parameters and the solution to the dual problem.

In the last chapter we will present and study concrete examples of the utility maximization portfolio problem in specific markets. First, we consider the complete markets case, where closed-form solutions are easily obtained. The detailed solution to the classical Merton problem with power utility function is provided. Lastly, we deal with incomplete markets under Itô processes and the Brownian filtration framework. The solution to the logarithmic utility function as well as to the power utility function is presented.

Opsomming

Die eerste benadering, begin deur Merton [Mer69, Mer71], om nutsmaksimering portefeulje probleme op te los in kontinue tyd is gebaseer op stogastiese beheerteorie. Merton se idee is om die maksimering portefeulje probleem te interpreteer as 'n stogastiese beheer probleem waar die handelstrategieë as 'n beheer-proses beskou word en die portefeulje waarde as die gereguleerde proses. Merton het die Hamilton-Jacobi-Bellman (HJB) vergelyking afgelei en vir die spesiale geval van die mags, logaritmies en eksponensiële nutsfunksies het hy 'n oplossing in geslote-vorm gevind. 'n Groot nadeel van hierdie benadering is die vereiste van die Markov eienskap vir die aandele pryse. Die sogenaamde *martingale metode* verteenwoordig die tweede benadering vir die oplossing van nutsmaksimering portefeulje probleme in kontinue tyd. Dit was voorgestel deur Pliska [Pli86], Cox en Huang [CH89, CH91] en Karatzas et al. [KLS87] in verskillende wisselvorme. Dit word aangevoer deur argumente van konvekse dualiteit, waar dit in staat stel om die aanvanklike dinamiese portefeulje optimalisering probleem te omvorm na 'n statiese een en dit op te los sonder dat' n "Markov" aanname gemaak hoef te word. 'n Bepalende antwoord (met die nodige en voldoende voorwaardes) tot die nutsmaksimering portefeulje probleem vir terminale vermoë is verkry deur Kramkov en Schachermayer [KS99]. In hierdie proefskrif bestudeer ons die konveks dualiteit benadering tot die verwagte nuts maksimering probleem (van terminale vermoë) in kontinue tyd stogastiese markte, wat soos reeds vermeld is teruggevoer kan word na die seminale werk van Merton [Mer69, Mer71]. Voordat ons die struktuur van ons tesis uitlê, wil ons graag beklemtoon dat die beginpunt van ons werk gebaseer is op Hoofstuk 7 van Pham [P09] se onlangse handboek. Die noukeurige leser sal egter opmerk, dat ons belangrike begrippe en resultate verdiep en bygelas het (soos die studie van die boonste (onderste) verskansing, die karakterisering van die noodsaaklike supremum van alle moontlike pryse, vergelyk Stelling 7.2.2 in Pham [P09] met ons verklaarde Stelling 2.4.9, die dinamiese programerings vergelyking 2.31, die superverskansing

stelling 2.6.1...) en ons het 'n aansienlike inspanning in die bewyse gemaak. Trouens, verskeie bewyse van stellings in Pham cite (P09) het ernstige gapings (nie te praat van setfoute nie) en selfs foute (kyk byvoorbeeld die bewys van Stelling 7.3.2 in Pham [P09] en ons bewys van Stelling 3.4.8). In die eerste hoofstuk, sit ons die verwagte nutsmaksimering probleem uit een en motiveer ons die konveks duaale benadering gebaseer op 'n voorbeeld van Rogers [KR07, R03]. Ons gee ook 'n kort oorsig van die von Neumann - Morgenstern Verwagte Nutsteorie. In die tweede hoofstuk, begin ons met die formulering van die superreplikasie probleem soos voorgestel deur El Karoui en Quenez [KQ95]. Die fundamentele resultaat in die literatuur oor super-verskansing is die duaale karakterisering van die versameling van alle eerste skenkings wat lei tot 'n super-verskans van' n Europese voorwaardelike eis. El Karoui en Quenez [KQ95] het eers die super-verskansing stelling 2.6.1 bewys in 'n Itô diffusie raamwerk en Delbaen en Schachermayer [DS95, DS98] het dit veralgemeen na, onderskeidelik, 'n plaaslik begrensde en onbegrensde semimartingale model, met 'n Hahn-Banach skeidings argument. Die superreplikasie probleem het 'n prag resultaat geïnspireer, genaamd die opsionele ontbinding stelling vir supermartingales 2.4.1 in stogastiese ontledings teorie. Hierdie belangrike stelling wat deur El Karoui en Quenez [KQ95] voorgestel is en tot volle veralgemening uitgebrei is deur Kramkov [Kra96] is uiteengesit in Afdeling 2.4 en bewys aan die einde van Afdeling 2.7. Die derde hoofstuk vorm die teoretiese basis van hierdie proefskrif en bevat die verklaring en gedetailleerde bewys van die beroemde Kramkov-Schachermayer stelling wat die dualiteit van nutsmaksimering portefeulje probleme adresseer. Eerstens, wys ons in Lemma 3.2.1 hoe om die dinamiese nutsmaksimering probleem te omskep in 'n statiese maksimerings probleem. Dit kan gedoen word te danke aan die duaale voorstelling van die versameling Europese voorwaardelike eise, wat oorheers (of super-verskans) kan word byna seker van 'n aanvanklike skenking x en 'n toelaatbare self-finansierings portefeulje strategie wat in Gevolgtrekking 2.5 gegee word en verkry is as gevolg van die opsionele ontbinding van supermartingale. In die tweede plek, met sekere aannames oor die nutsfunksie, is die bestaan en uniekheid van die oplossing van die statiese probleem gegee in Stelling 3.2.3. Omdat die oplossing van die statiese probleem nie maklik verkrygbaar is nie, sal ons kyk na die duaale vorm. Ons sintetiseer dan die duale probleem van die primêre probleem met konvekse toegevoegde funksies. Voordat ons die Kramkov-Schachermayer Stelling 3.4.1 beskryf, gee ons die Inada voorwaardes en die Asimptotiese Elastisiteits Voorwaarde vir Nutsfunksies. Ter wille van duidelikheid, verdeel ons die lang en tegniese bewys van die Kramkov-Schachermayer Stelling ref in verskeie lemmas en proposisies op, elk van onafhanklike belang waar die nodige aannames duidelik uiteengesit is vir elke stap van die bewys. Die belangrikste argument in die bewys van die Kramkov-Schachermayer Stelling is 'n oneindig-dimensionele weergawe van die minimax stelling (die klassieke metode om 'n saalpunt vir die Lagrange-funksie te bekom is nie genoeg in die geval nie), wat noodsaaklik is in die teorie van Lagrange-multiplikators. Vir die, meld en bewys ons die tegniese Lemmata 3.4.5 en 3.4.6. Die belangrikste stappe in die bewys van die die Kramkov-Schachermayer Stelling 3.4.1 is:

- Ons wys in Proposisie 3.4.9 dat die oplossing vir die duale probleem bestaan en ons karaktiriseer dit in Proposisie 3.4.12.
- Uit die konstruksie van die duale probleem vind ons 'n versameling nodige en voldoende voorwaardes (3.1.1), (3.1.2), (3.3.1) en (3.3.7) wat die primêre en duale probleem oplossings elk moet aan voldoen.
- Deur hierdie voorwaardes te gebruik, kan ons die bestaan van die oplossing vir die gegewe probleem wys en dit karakteriseer in terme van die mark parameters en die oplossing vir die duale probleem.

In die laaste hoofstuk sal ons konkrete voorbeelde van die nutsmaksimering portefeulje probleem bestudeer vir spesifieke markte. Ons kyk eers na die volledige markte geval waar geslote-vorm oplossings maklik verkrygbaar is. Die gedetailleerde oplossing vir die klassieke Merton probleem met mags nutsfunksie word voorsien. Ten slotte, hanteer ons onvolledige markte onderhewig aan Itô prosesse en die Brown filtrering raamwerk. Die oplossing vir die logaritmiese nutsfunksie, sowel as die mags nutsfunksie word aangebied.

Acknowledgements

This thesis would not have seen the day without the support of the entire AIMS family. First, I thank Dr. Ghomrasni my supervisor for his encouragement and restless effort he put in during the preparation of this thesis. I owe many thanks to Prof. Fritz Hahne (former AIMS director) and Prof. Barry Green the director of AIMS for giving me the opportunity to further my studies in sciences and specially in mathematical finance. I also thank Prof. Frittelli for exciting discussions on the utility maximization portfolio problems and duality approach during the 3rd summer school in mathematical finance at AIMS and his interest in my work. I also wish to thank Ms. Frances Aron for revising the English of my manuscript. I would like to express my sincere gratitude to AIMS for funding my masters thesis. I am thankful to Dr. Lafras Uys for the Afrikaans translation of the abstract of this thesis.

Contents

Chapter 1

Introduction

Nowadays, financial theory is one of the major economic fields where decision-making under uncertainty plays a crucial part. For example, the problem of maximizing the expected utility of an economic agent who invests in a financial market. In the frame-work of a continuous-time complete financial model, the problem was examined for the first time by R. Merton in two seminal papers [Mer69, Mer71]. He derived the Bellman equation for the value function of the optimization problem by using Itô calculus and the method of stochastic control theory. However, this method requires the Markov property for the state processes. The martingale and convex duality method, as an alternative approach to the problem, allows us to work in non-Markovian settings. In the complete markets case, this methodology was devised by Pliska [Pli86], Cox and Huang [CH89] and Karatzas, Lehoczky and Shreve [KLS87] by providing powerful insights into the solutions of such problems to prove the form of the optimal solution to significant generalizations of the original Merton [Mer69] problem. In the incomplete markets framework, the problem was studied by He and Pearson [HP91a, HP91b] and Karatzas et al. [KLSX91] for some specified model.

There is, actually, a unified and easy approach to finding the dual form of the problem, which works in a varied range of situations. It may be seen as the Pontryagin approach to dynamic programming; or interpreted in the Hamiltonian language of Bismut [BI73, BI75]. Before moving onto the heart of our topic, let us present the method applied to the simplest example to illustrate our motivation.

1.1 Introduction to Optimal Investment

Consider an economic agent (an investor) in an *arbitrage-free* financial model, with initial capital x and her goal is to invest x "optimally" up to maturity T. A natural question is: how to compare two investment strategies:

1. $x \longrightarrow X_T = X_T(\omega)$
2. $x \longrightarrow Y_T = Y_T(\omega)$

?. Clearly, we would prefer the first to the second if

$$X_T(\omega) \geq Y_T(\omega)\,, \quad \forall\, \omega \in \Omega.$$

However, as the model is arbitrage-free, if this inequality holds, we must have

$$X_T(\omega) = Y_T(\omega)\,, \quad \omega \in \Omega.$$

The classical approach (Von Neumann - Morgenstern, Savage) is that the investor is "quantified" by P, a "scenario" probability measure and a utility function $U = U(x)$. The quality of a strategy $x \longrightarrow X_T = X_T(\omega)$ is then measured by expected utility $E[U(X_T)]$. Given two strategies $x \longrightarrow X_T$ and $x \longrightarrow Y_T$ the investor will prefer the first one if

$$E[U(X_T)] \geq E[U(Y_T)]\,.$$

Therefore, our problem is to find an optimal investment strategy $x \longrightarrow \hat{X}_T$ such that

$$E[U(\hat{X}_T)] = u(x) = \sup_{X \in \Xi(x)} E[U(X_T)]$$

1.2 Motivation

The following example is taken from Rogers [KR07, R03]. Let us consider an investor who may trade in any of $n \geq 1$ risky assets $S = (S^1, \cdots, S^n)$ with dynamics given by $dS_t = S_t(\sigma_t dW_t + b_t dt)$ and in a riskless bank account S^0 with dynamics $dS^0_t = r_t S^0_t dt$ generating interest at rate r_t. It can be easily seen that the dynamics of the investor wealth process X corresponding to a self-financing portfolio strategy (without consumption) is given by

$$dX_t = r_t X_t dt + \theta_t(\sigma_t dW_t + (b_t - r_t \mathbf{1})dt), \qquad X_0 = x, \tag{1.1}$$

where all processes are adapted to the flow of information (or "filtration") generated by the driving standard d-dimensional Brownian motion W, the volatility σ is a $n \times d$ matrix-valued process, and all other processes have the dimensions implied by (1.1). For concreteness, we are assuming that negative wealth is not allowed, in other words X must remain nonnegative, i.e. $X_t \geq 0$, $\forall t \geq 0$. The process $\theta = (\theta^1, \theta^2, \cdots, \theta^n)$ is the n-dimensional vector of amounts of wealth invested in each of the stocks $(S^i)_{0 \leq i \leq n}$. The investor aims to maximize his wealth portfolio at the end of the investment period $[0, T]$, where $T > 0$ is a fixed finite time-horizon, i.e. to find

$$\sup_{\theta} E[U(X_T)], \tag{1.2}$$

where the utility function $U(\cdot)$ (see Section 1.3 for more information) satisfies the Inada conditions (see Section 3.3.1 for details).

The dynamics (1.1) of X must satisfy various constraints such as the *bankruptcy constraint* (i.e. choice of θ such that $X \geq 0$), and we transform the constrained optimization problem (1.2) into an *unconstrained* optimization problem by introducing appropriate Lagrange multipliers (see Section 1.4 for more information about the Lagrange multipliers method). In this section, we deal only with the constraint $X \geq 0$ and the corresponding problem

$$\sup_{\{\theta \,|\, X \geq 0\}} E[U(X_T)]\,. \tag{1.3}$$

To this end, consider the positive process Y satisfying the following dynamics

$$dY_t = Y_t(\beta_t dW_t + \alpha_t dt), \quad \text{with} \quad Y_0 > 0\,, \tag{1.4}$$

and let us evaluate the stochastic integral $\int_0^T Y_s dX_s$. On the one hand, integration by parts formula gives immediately

$$\int_0^T Y_s dX_s = X_T Y_T - X_0 Y_0 - \int_0^T X_s dY_s - \langle X, Y \rangle_T, \tag{1.5}$$

and on the other hand using (1.1) leads to

$$\int_0^T Y_s dX_s = \int_0^T Y_s \theta_s \sigma_s dW_s + \int_0^T Y_s \{r_s X_s + \theta_s(b_s - r_s \mathbf{1})\} ds. \tag{1.6}$$

Using the fact that the covariation $\langle X, Y \rangle_T$ at time T of X and Y is easily computed and given by

$$\langle X, Y \rangle_T = \int_0^T Y_s \beta_s \theta_s \sigma_s ds\,, \tag{1.7}$$

and suppose that expectations of stochastic integrals with respect to the Brownian motion W vanish, the expectation of $\int_0^T Y_s dX_s$ is from (1.5)

$$E[X_T Y_T - X_0 Y_0 - \int_0^T Y_s\{\alpha_s X_s + \theta_s \sigma_s \beta_s\} ds], \tag{1.8}$$

and from (1.6)

$$E[\int_0^T Y_s\{r_s X_s + \theta_s(b_s - r_s\mathbf{1})\} ds], \tag{1.9}$$

Since (1.8) and (1.9) must be equal for any feasible X, we obtain the condition that the Lagrangian function

$$\begin{aligned} \Lambda(Y) \equiv & \sup_{X\geq 0,\theta} E[U(X_T) + \int_0^T Y_s\{r_s X_s + \theta_s(b_s - r_s\mathbf{1})\} ds \\ - & X_T Y_T + X_0 Y_0 + \int_0^T Y_s\{\alpha_s X_s + \theta_s \sigma_s \beta_s\} ds], \qquad (1.10) \\ = & \sup_{X\geq 0,\theta} E[U(X_T) - X_T Y_T + X_0 Y_0 \\ + & \int_0^T Y_s\{r_s X_s + \theta_s(b_s - r_s\mathbf{1}) + (\alpha_s X_s + \theta_s \sigma_s \beta_s)\} ds] \qquad (1.11) \end{aligned}$$

is an upper bound for the value in (1.3) whatever the choice of Y we consider, and will hopefully be equal to it if we minimize over Y.

In the definition of $\Lambda(Y)$ we require that $X_T \geq 0$ and that $X_s \geq 0\,(0 \leq s < T)$. Now the maximization of 1.11 over $X_T \geq 0$ is very easy - we obtain

$$\begin{aligned} \Lambda(Y) = & \sup_{X\geq 0,\theta} E[\widetilde{U}(Y_T) + X_0 Y_0 \qquad (1.12) \\ + & \int_0^T Y_s\{r_s X_s + \theta_s(b_s - r_s\mathbf{1}) + (\alpha_s X_s + \theta_s \sigma_s \beta_s)\} ds] \qquad (1.13) \end{aligned}$$

where $\widetilde{U}(y) = \sup_x (U(x) - xy)$ is the Legendre-Fenchel transformation (or convex dual) of U. The maximization over $X_s \geq 0$ results in a finite value if, and only if, the complementary slackness condition

$$r_s + \alpha_s \leq 0 \tag{1.14}$$

holds, and maximization over $\theta_s \in \mathbb{R}$ results in a finite value if, and only if, the complementary slackness condition

$$\sigma_s \beta_s + b_s - r_s\mathbf{1} = 0 \tag{1.15}$$

holds. We therefore add these constraints. The maximized value is then

$$\Lambda(Y) = E[\widetilde{U}(Y_T) + X_0 Y_0]. \tag{1.16}$$

The dual problem therefore ought to be

$$\inf_Y \Lambda(Y) = \inf_Y E[\widetilde{U}(Y_T) + X_0 Y_0], \tag{1.17}$$

with Y defined by (1.4), where α and β are understood to satisfy the complementary slackness conditions (1.14) and (1.15). Actually, because the convex conjugate $\widetilde{U}(\cdot)$ is a decreasing function, a little thought shows that we want Y to be big, so that the drift or the "discount rate" α will be as large as it can be, that is, the inequality (1.14) will actually hold with equality i.e. we should have $\alpha_t = -r_t$.

We can interpret the multiplier process Y which satisfies the dynamics $dY_t = Y_t dN_t$ where $dN_t = \beta_t dW_t - r_t dt$, now written as a Doléans exponential

$$Y_t = Y_0 \mathcal{E}(N)_t = Y_0 \exp\{\int_0^t \beta_s dW_s - \int_0^t r_s ds - \frac{1}{2}\int_0^t \beta_s^2 ds\} = Y_0 \exp\{-\int_0^t r_s ds\} \cdot Z_t\,,$$

as the product of the initial value Y_0, the riskless discounting term $\exp\{-\int_0^t r_s ds\}$, and a (change-of-measure) martingale Z, with

$$Z_t = \mathcal{E}(\beta dW)_t = \exp\{\int_0^t \beta_s dW_s - \frac{1}{2}\int_0^t \beta_s^2 ds\}\,, \quad \text{for} \quad t \in [0,T],$$

where the process β satisfies $\sigma_s \beta_s = -(b_s - r_s \mathbf{1})$ where the LHS is minus the risk-premium process λ (see Section 2.7 for more details), in other words, its effect is to convert the rates of return of all stocks into the riskless rate. In conclusion, we have the multiplier process Y in the form of

$$Y_t = Y_0 \exp\{-\int_0^t r_s ds\} \cdot \exp\{\int_0^t -\lambda_s dW_s - \frac{1}{2}\int_0^t \lambda_s^2 ds\}\,, \quad \text{for} \quad t \in [0,T],$$

In the complete market case with $n = d$ and σ having bounded inverse, we recover (see Subsection 3.3.2 for details) the well-known result of Karatzas et al. [KLS87], given by

$$U'(X_T^\star) = Y_T\,, \tag{1.18}$$

$$\text{with} \quad E[Y_T X_T^\star] = x\,. \tag{1.19}$$

In other words the marginal utility U' of terminal optimal wealth $X^\star$ is the pricing kernel, or state price density Y_T.

As we can see, our motivation above required the knowledge of the notions of utility function and Lagrange multiplier. Thus, in the next sections, we will introduce these notions in order to familiarize the readers with them.

1.3 Utility Function

In order to model any decision problem under risk, it is necessary to introduce a functional representation of preferences which measures the degree of satisfaction of the decision maker. Basically, the purpose of the utility functions is to allow us to see preference relations among various levels of consumption, various strategies for asset holdings, etc. The investor is supposed to be rational: this means that his choices are made according to given good rules which are stable over time (in some sense). Thus a binary relation on possible outcomes can be proposed to analyze his behavior. Specific axioms (see e.g. [Pr07]) are introduced to describe his rationality. Then, for this given identified choice functional, his optimal decision (for example his investment strategy) is determined from the maximization of this criterion.

1.3.1 The von Neumann - Morgenstern Expected Utility Theory

The theory of von Neumann-Morgenstern provides a numerical representation of an individual's preferences over lotteries for the case of choice under uncertainty. Mathematically, a lottery is a probability distribution defined on the set of payoffs and it can be discrete, continuous and mixed. For a recent account of von Neumann-Morgenstern theory we refer the reader to ([RSF08]). In the continuous case, for a random payoff X, the lottery is equivalently described by the probability distribution P_X or by the cumulative distribution function (c.d.f.)* F_X. For example, any portfolio of assets with payoff X (at a given and fixed time) may be seen as a continuous lottery. Let denote by $\mathcal{L}$ the set of all lotteries. Any element of $\mathcal{L}$ is considered a possible choice of an economic agent. For $P_X, P_Y \in \mathcal{L}$ we need only consider three cases:

- The economic agent may prefer P_X to P_Y or there is no clear preference between the two, denoted by $P_X \succeq P_Y$.
- The economic agent may prefer P_Y to P_X or there is no clear preference between the two, denoted by $P_Y \succeq P_X$.
- If both relations hold, $P_Y \succeq P_X$ and $P_X \succeq P_Y$, then the economic agent is said to be indifferent between the two choices $P_X \sim P_Y$.

*Recall that the c.d.f. $F_X(x)$ is the probability that the payoff is below x, i.e. $P(X \leq x) = F_X(x)$.

A numerical representation of a preference order is a real-valued function U defined on the set of lotteries, $U\colon \mathcal{L} \to \mathbb{R}$, with the property that $P_X \succeq P_Y$ if and only if $U(P_X) \geq U(P_Y)$. Note that such a numerical representation of a preference order is not unique.

The von Neumann-Morgenstern theory asserts that if the preference order is subject to certain technical continuity conditions, then the numerical representation U has the form

$$U(P_X) = \int_{\mathbb{R}} u(x) dF_X(x), \tag{1.20}$$

where $u(\cdot)$ is the utility function of the economic agent defined over the elementary outcomes of the random variable X with c.d.f. F_X. Equality (1.20) is in fact the mathematical expectation of the random variable $u(X)$, i.e. we have

$$U(P_X) = E[u(X)], \tag{1.21}$$

and therefore the numerical representation of the preference order is the expected utility.

Remark 1.3.1 *If the lottery is discrete and finite, then the payoff is a discrete finite random variable and equation (1.20) becomes*

$$U(P_X) = \sum_{j=1}^{j=n} u(x_j) p_j, \tag{1.22}$$

where x_j denote the outcomes and p_j is the probability that the $j-th$ outcome occurs, $p_j = P(X = x_j)$.

1.3.2 Types of Utility Functions

Generally, the utility function properties characterize the investors preferences. For example, the utility functions need to have the desired property of being *non-decreasing* (any investor is insatiable, she prefer more to less) i.e. we have

$$u(x) \leq u(y), \qquad \text{if} \qquad x \leq y \qquad \text{for any} \qquad x, y \in \mathbb{R}.$$

The outcomes x and y can be interpreted as the payoffs of two opportunities without an element of uncertainty, which means that both x and y occur with probability one.

For the *risk averse* investor the utility function is *concave*. Indeed, assume that the payoff has two possible outcomes, x_1 with probability $p \in [0, 1]$ and x_2 with probability

$1-p$. The expected payoff is equals $px_1+(1-p)x_2$. The risk-aversion property is expressed in term of the utility function as

$$u(px_1 + (1-p)x_2) \geq pu(x_1) + (1-p)u(x_2) \qquad \forall x_1, x_2 \quad \text{and} \quad p \in [0,1]\,, \tag{1.23}$$

where the LHS corresponds to the utility of the certain payoff $px_1 + (1-p)x_2$ and the RHS is the expected utility of the payoff. An *absolute risk aversion* is measured by the coefficient of absolute risk aversion (CARA) defined by

$$r_A(x) = -\frac{u''(x)}{u'(x)}, \tag{1.24}$$

which shows that the more curved the utility function is, the higher the risk-aversion level of the investor.

In the rest of this section we describe some common utility functions

1. A linear utility function is defined by

$$u(x) = a + bx.$$

   The linear utility function satisfies (1.23) with equality and represents a risk-neutral (indifferent to risk) investor. Moreover, when $b > 0$, it represents a insatiable investor.

2. A quadratic utility function is defined by

$$u(x) = a + bx + cx^2.$$

   The quadratic utility function is concave for $c < 0$ and in this case represents a risk-averse investor.

3. A logarithmic utility function is defined by

$$u(x) = \ln(x), \qquad x > 0.$$

   It represents a insatiable, risk-averse investor. The CARA is given $r_A(x) = 1/x$ (notice that it decreases with x).

4. Exponential utility function is defined by

$$u(x) = -e^{-ax}, \qquad a > 0.$$

   It represents a insatiable, risk-averse investor with a constant CARA $r_A(x) = a$.

5. Power utility function is defined by

$$u(x) = -\frac{x^{-a}}{a}, \qquad a > 0, x > 0.$$

It represents an insatiable, risk-averse investor with a decreasing CARA $r_A(x) = \frac{a+1}{x}$.

1.4 Motivation of the Lagrange Formulation

Our aim is to minimize a function subject to a constraint. More precisely, let us consider a differentiable function $F : \mathbb{R}^d \times \mathbb{R}^n \longrightarrow \mathbb{R}$. Our goal is to find

$$\begin{cases} \min_{x \in A} F(x, y), \\ \text{subject to } \; y = g(x)\,, \end{cases} \tag{1.25}$$

for a given differentiable function $g : \mathbb{R}^d \longrightarrow \mathbb{R}^n$ and a compact set $A \subset \mathbb{R}^d$. The problem (1.25) yields to the classical condition for an interior minimum

$$\frac{d}{dx} F(x, g(x)) = \partial_x F(x, g(x)) + \partial_x g(x) \partial_y F(x, g(x)) = 0 \tag{1.26}$$

One of the best method to solve problem (1.25) is to write the Lagrangian function $L(\lambda, y, x) = F(x, y) + \lambda.(y - g(x))$, where $\lambda \in \mathbb{R}^n$ is the Lagrange multiplier. Suppose that we minimize F with respect to all three variables. Then the usual necessary condition for an interior minimum is as follows:

$$\partial_\lambda L(\lambda, y, x) \;=\; y - g(x) = 0, \tag{1.27}$$

$$\partial_y L(\lambda, y, x) \;=\; \partial_y F(x, y) + \lambda = 0 \tag{1.28}$$

$$\partial_x L(\lambda, y, x) \;=\; \partial_x F(x, y) - \lambda \partial_x g(x) = 0. \tag{1.29}$$

Observe that Equation (1.27) is precisely the required constraint $y = g(x)$. From Equation (1.28) we derive the Lagrange multiplier to be $\lambda = -\partial_y F(x, y)$. Lastly, Equation (1.28) implies that for this choice of Lagrange multiplier we have $\partial_x L(-\partial_y F(x, y), y, x) = \frac{d}{dx} F(x, g(x))$. In other words, the Lagrange multiplier is chosen exactly such that the partial derivative with respect to x of the Lagrangian function equals the total derivative of the objective function $F(x, g(x))$ to be minimized.

Remark 1.4.1 *In general, we use the Lagrange principle when the constraint is given implicitly. For example, as* $f(x,y) = 0$ *with a differentiable* $f : \mathbb{R}^d \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$*. In this case, the required condition* $\det \partial_y f(x,y) \neq 0$ *in the implicit function theorem yields that the function* $y(x)$ *is well defined and satisfies* $f(x, y(x)) = 0$ *with* $\partial_x y = -\partial_y f(x,y)^{-1} \partial_x f(x,y)$*, thus the Lagrange multiplier method works as well.*

Chapter 2

Dual representation of the Superreplication Cost

In this chapter we formulate the superreplication problem and we present the optional decomposition of supermartingale theorem which plays a fundamental role in the dual characterization of the superreplication cost. The delicate proof of the optional decomposition of super-martingale theorem is given in the Itô processes and Brownian filtration framework. This result will be crucial in the next chapter for establishing the equivalent formulation between the dynamic optimization problem and the static one.

2.1 Introduction

The basic idea of martingale methods in portfolio optimization problems is to reduce the initial *dynamic* problem, which consists of an optimization over a control process, to an optimization problem on the state variable given by the terminal value of the portfolio (i.e. *static*) with a linear constraint described as a change of an equivalent probability measure by a Radon-Nikodým density called in this context a dual variable. Let us begin by illustrating this idea in a simple example borrowed from Pham ([P09]). Consider a state process $X = (X_t)_{0\leq t\leq T}$, controlled by a progressively measurable process $\alpha = (\alpha_t)_{0\leq t\leq T}$, with dynamics given by

$$dX_t = \alpha_t(dt + dW_t)\,, \quad 0 \leq t \leq T,$$

where W is a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. We assume that the filtration (or "flow of information") $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ is the natural filtration generated by the driving Brownian motion W. For a positive real number x and a control process α, a (strong) solution to the above SDE with initial condition $X_0^x = x$ is denoted by X^x and let $\mathcal{A}(x)$ be the set of control processes α such that $X_t^x \ge 0$ for all $t \in [0, T]$, i.e. the state process $X = (X_t)_{0\le t\le T}$ remains nonnegative. Given a utility function (i.e. a concave and increasing function) U on $\mathbb{R}_+$, the expected utility optimization problem is given by

$$v(x) = \sup_{\alpha\in\mathcal{A}(x)} E[U(X_T^x)], \quad x \ge 0. \tag{2.1}$$

Using Girsanov theorem, we can obtain an equivalent probability measure $Q \sim P$, under which the process $B = (B_t = W_t + t)_{0\le t\le T}$ a standard Brownian motion. Let $L^0_+(\Omega, \mathcal{F}_T, P)$ be the space of nonnegative $\mathcal{F}_T$-measurable random variables. For any r.v. $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$ satisfying the constraint $E^Q[X_T] \le x$, there exists, as an application of the Itô representation Theorem C.0.5 under Q, a process $\alpha \in \mathcal{A}(x)$, such that

$$X_T = E^Q[X_T] + \int_0^T \alpha_t dB_t \le X_T^x = x + \int_0^T \alpha_t dB_t. \tag{2.2}$$

Conversely, for any control process $\alpha \in \mathcal{A}(x)$, the controlled process $X^x = x + \int \alpha dB = x + \int \alpha(dt + dW)$ is a non-negative local martingale under the probability Q, hence, a Q supermartingale, and thus we have $E^Q[X_T] \le x = E^Q[X_0^x]$. This shows that

$$\{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) \mid E^Q[X_T] \le x\} = \{X_T^x \mid \alpha \in \mathcal{A}(x)\}.$$

Therefore the optimisation problem (2.1) can be stated in an equivalent way as

$$\begin{cases} v(x) = \sup\limits_{X_T\in L^0_+(\Omega,\mathcal{F}_T,P)} E[U(X_T)], \\ \text{subject to } E[\frac{dQ}{dP}X_T] \le x \end{cases} \tag{2.3}$$

We are then left with a concave optimization problem in the infinite dimension space $L^0_+(\Omega, \mathcal{F}_T, P)$ subject to a linear constraint represented by the Radon-Nykodým density dQ/dP as a dual variable. Therefore, the classical convex analysis techniques may now be applied for solving problem (2.3).

The Itô representation Theorem, valid under a Brownian filtration, played a central role in the above equivalent dual resolution approach. In order to tackle more general expected utility optimization problems, notably when the equivalent probability measure Q is not unique (in incomplete markets there will be an infinite of them), we will need the optional decomposition for supermartingales theorem, which is a deep result in stochastic analysis.

The optional decomposition for supermartingales theorem, first established by El Karoui and Quenez [KQ95] in the framework of Itô diffusion processes, was initially motivated by an important problem in mathematical finance the so-called superreplication problem in incomplete markets, and will be discussed in more details in Section 2.2. This important theorem has been subsequently extended by Kramkov and coauthors [Kra96, FK97] to the general framework of (not necessarily continuous) semimartingales processes. This result will be presented in Section 2.4 below and a proof will be provided under the Brownian filtration framework in 2.7.5.

2.2 Formulation of the Superreplication Problem

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a complete filtered probability space satisfying the usual conditions. The real $T > 0$ is a fixed finite horizon $T < \infty$, but we remark that the results described in this thesis can also be extended to the case of an infinite horizon. For simplicity, we assume that $\mathcal{F}_0$ is trivial, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and also that $\mathcal{F} = \mathcal{F}_T$.

We consider a financial market which consists of one risk-free asset and n stocks. Without loss of generality, we will always consider the price process of the risk-free asset to be constant equal to 1 (because we always may choose the risk-free asset as numéraire). The (discounted) price process $S = (S^i)_{1\leq i\leq n}$ of the n stocks is assumed to be a continuous $\mathbb{R}^n$-valued semimartingale, on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, P)$.

Let $L(S)$ be the set of progressively measurable processes α, integrable with respect to S (i.e. $\int |\alpha|\cdot dS < \infty$). An element $\alpha = (\alpha^i)_{1\leq i\leq n} \in L(S)$ represents a portfolio strategy for an investor: α^i_t is the (real) number (when $\alpha^i_t > 0$ you are long, when $\alpha^i_t < 0$ you are short) of shares invested in the stock S^i at time t. Thus, starting at time $t = 0$ with an initial capital $x \in \mathbb{R}$, the (discounted) wealth process of the investor following the (self-financing) portfolio strategy α is

$$x + \int_0^t \alpha_s \cdot dS_s = x + \sum_{i=1}^n \int_0^t \alpha^i_s \cdot dS^i_s, \qquad 0 \leq t \leq T. \tag{2.4}$$

Since we work with continuous-time trading strategies, we need to eliminate suicide strategies as well as doubling strategies by adding constraints on the set of self-financing trading strategies. A control process $\alpha \in L(S)$ is said to be admissible if $\int \alpha dS$ is lower-bounded (i.e. for an $\alpha \in L(S)$ there exists a constant c_α with the property that $\int_0^t \alpha_s \cdot dS_s \geq c_\alpha$ for $t \in [0,T]$), and we denote by $\mathcal{A}(S)$ the set of such admissible controls. This admissibility prevent doubling strategies (for more details we refer to Harrison and Pliska [HP81]): since otherwise, one could construct (even in an arbitrage-free and complete financial market) a sequence of portfolio strategies $(\alpha^n)_{n\geq 1}$ with the property that $\int_0^T \alpha_s^n dS_s \to \infty$ a.s., which represents a means to earn as much money as desired at time T from a zero initial endowment!.

Let X_T be a contingent claim of maturity T, that is, a nonnegative, $\mathcal{F}_T$-measurable random variable. In other words, we consider a European-type option X_T, whose payoff is made at the terminal (maturity) date T and may depend on the whole history up to T. The superreplication problem of the contingent claim X_T consists in finding the minimal initial capital that allows us to dominate (or superhedge) in the almost sure sense the contingent claim at maturity. Mathematically, this problem is stated as

$$v_0(X_T) = \inf\left\{x \in \mathbb{R} : \exists \alpha \in \mathcal{A}(S),\ \ x + \int_0^T \alpha_t dS_t \geq X_T \text{ a.s.}\right\}, \tag{2.5}$$

with the convention that $\inf \emptyset = \infty$. $v_0(X_T)$ is called the superreplication cost or super-hedging price of X_T, and if $v_0(X_T)$ attains the infimum in (2.5), the control $\alpha \in \mathcal{A}(S)$ with the property that $v_0(X_T) + \int_0^T \alpha_t dS_t \geq X_T$ is called the superreplication portfolio strategy. A contingent claim X_T is called attainable if there exists a superreplication portfolio strategy $\alpha \in \mathcal{A}(S)$ such that $X_T = v_0(X_T) + \int_0^T \alpha_t dS_t$, i.e. we have equality.

The super-hedging price $v_0(X_T)$ is the smallest amount of initial capital which allows to eliminate all shortfall risk. However, if the option is not attainable (this may happen in incomplete markets), the super-hedging price allows for arbitrage. Hence the super-hedging price $v_0(X_T)$ must exceed the option-premium (fair price) in an arbitrage-free market.

The fundamental result in the literature on super-hedging is the dual characterization of the set $\mathcal{D}^{X_T}$ of all initial endowments $x \in \mathbb{R}$ leading to super-hedge X_T, i.e.

$$D^{X_T} := \left\{x \in \mathbb{R} : \exists \alpha \in \mathcal{A}(S),\ \ x + \int_0^T \alpha_t dS_t \geq X_T \text{ a.s.}\right\}. \tag{2.6}$$

Of course, if not empty, it is a semi-infinite interval (possibly, coinciding with the whole real line $\mathbb{R}$). A priori, it can be either closed or open, i.e. of the form $[\bar{x}, \infty)$ or $(\bar{x}, \infty)$ with $\bar{x} = v_0(X_T)$.

Similarly, we can define the class of lower-hedges of X_T, by

$$-D^{-X_T} := \left\{ x \in \mathbb{R} : \exists \alpha \in \mathcal{A}(S), \;\; -x + \int_0^T \alpha_t dS_t \geq -X_T \text{ a.s.} \right\} . \tag{2.7}$$

It can be either of the form $(-\infty, \underline{x})$ or $(-\infty, \underline{x}]$. The "fair" prices of X_T lie in the interval $[\underline{x}, \bar{x}]$.

In an incomplete frictionless market, the relevant dual variables are the densities of all equivalent martingale measures dQ/dP. We will denote by $\mathcal{M}_e(S)$ the set of all equivalent (local) martingale measures for S. In this setting, the superhedging theorem 2.6.1 states that

$$D^{X_T} = \left\{ x \in \mathbb{R} \,:\, x \geq E^Q[X_T] \,, \forall Q \in \mathcal{M}_e(S) \right\} . \tag{2.8}$$

We note that the following inclusion

$$D^{X_T} \subseteq \left\{ x \in \mathbb{R} \,:\, x \geq E^Q[X_T] \,, \forall Q \in \mathcal{M}_e(S) \right\} . \tag{2.9}$$

is obvious. To show the opposite inclusion, we need to apply a fundamental result known as the optional decomposition theorem. This will be done in Section 2.6.

An important consequence of (2.44) is that the super-hedging price $v_0(X_T)$ satisfies

$$v_0(X_T) = \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \,, \tag{2.10}$$

and we have $D^{X_T} = [v_0(X_T), \infty)$.

In an arbitrage-free and complete market, the (super-)hedging price $v_0(X_T)$ at time $t = 0$ of a contingent claim X_T, coincides with the expectation of (discounted) X_T under the unique equivalent martingale measure Q, i.e. $v_0(X_T) = E^Q[X_T]$. When the context is clear, we will simply write v_0 instead of $v_0(X_T)$.

While an advantage of super-hedging is that it is preference-free, from the previous characterization of v_0 as the biggest expectation $E^Q[X_T]$ over all equivalent martingale measures, it becomes apparent that pursuing a super-hedging strategy can be too expensive, depending on the financial model and on the constraints on portfolios. This is the main disadvantage of such a criterion, which is nonetheless of great interest as a benchmark.

El Karoui and Quenez [KQ95] first proved the superhedging theorem in an Itô's diffusion setting and Delbaen and Schachermayer [DS95, DS98] a generalized it to, respectively, a locally bounded and unbounded semimartingale model, using a Hahn-Banach separation argument.

The super-hedging theorem can be extended in order to characterize the dynamics of the minimal super-hedging portfolio of a contingent claim X_T, i.e. the cheapest at any time t of all superhedging portfolios of $E^Q[X_T]$ with same initial wealth. This extension is a consequence of the so-called optional decomposition of super-martingales and will be discussed in Section 2.4.

For any $x \in \mathbb{R}_+$, we define the set

$$\mathcal{C}(x) = \left\{ X_T \in L^0_+(\Omega, \mathcal{F}_T, P) : \exists \alpha \in \mathcal{A}(S), x + \int_0^T \alpha_t dS_t \geq X_T \text{ a.s.} \right\}, \tag{2.11}$$

In other words, $\mathcal{C}(x)$ represents the set of European contingent claims, which can be dominated almost surely from an initial capital x and an admissible portfolio strategy $\alpha \in \mathcal{A}(S)$.

The purpose of this Chapter is to present a probabilistic representation and characterization of both the super-hedging price v_0 and the set $\mathcal{C}(x)$, which will prove to be extremely useful in Chapter 3, in terms of some dual space of probability measures.

2.3 Equivalent Martingale Measures and no Arbitrage Principle

We define

$$\mathcal{M}_e(S) := \{Q \sim P \text{ on } (\Omega, \mathcal{F}_T) : S \text{ is a } Q-\text{local martingale}\}. \tag{2.12}$$

$\mathcal{M}_e(S)$ is called the set of equivalent local martingale measures (in short E(local)MM) or risk-neutral probability measures. In other words, $Q \in \mathcal{M}_e(S)$ is called an equivalent local martingale measure if Q is equivalent to P and it is denoted by $Q \sim P$, together with the fact that S is a local martingale under Q.

Throughout this thesis, we make the crucial standing assumption that the set of equivalent local martingale measures $\mathcal{M}_e(S)$ is noempty:

$$\mathcal{M}_e(S) \neq \emptyset. \tag{2.13}$$

Since we are dealing with *continuous* semimartingales S only (in which case the notion of *sigma-martingale* coincides with local martingale), this mathematical assumption is equivalent to the no free lunch (with vanishing risk) condition, which is a refinement of the no arbitrage condition which is of a paramount importance in mathematical finance, and we refer the interested reader to the seminal papers by Delbaen and Schachermayer [DS94, DS95, DS98, DS99] for this result, known as the first fundamental theorem of asset pricing. We add here a fact, which we will be often used throughout this thesis: for any E(local)MM $Q \in \mathcal{M}_e(S)$ and admissible process $\alpha \in \mathcal{A}(S)$, the lower-bounded stochastic integral $\int \alpha dS$ is by definition a Q-local martingale, therefore a Q-supermartingale as a consequence of Fatou lemma.

We then obtain $E^Q[\int_0^T \alpha_t dS_t] \leq 0 = E^Q[\int_0^0 \alpha_t dS_t]$. Therefore, condition (2.13) implies

$$\nexists\, \alpha \in \mathcal{A}(S)\,, \quad \int_0^T \alpha_t dS_t \geq 0\,, \ \ a.s. \quad \text{and} \quad P\big(\int_0^T \alpha_t dS_t > 0\big) > 0\,.$$

meaning that one cannot find an admissible self-financing portfolio strategy, which allows us, starting from a null capital, to reach almost surely at T a nonnegative wealth, with a nonzero probability of being strictly positive. This is the economical condition of no arbitrage.

2.4 Optional Decomposition of Super-Martingale Theorem

The superreplication problem inspired a very nice result, called the optional decomposition theorem for supermartingales, in stochastic analysis theory, which we state in the general continuous semimartingale case (since we have chosen S to be continuous semimartingales). This is a very deep result of general theory of stochastic processes. The optional decomposition was first proved by El Karoui and Quenez in [13, 14] for diffusions and then extended to general semimartingales by Kramkov [24], Föllmer and Kabanov [15] and Delbaen and Schachermayer [12]. A complete proof of the optional decomposition theorem for supermartingales, under the Itô processes framework, is also presented at the end of this chapter.

Theorem 2.4.1 ***(Optional decomposition of supermartingale theorem)***

Let X be a nonnegative càd-làg process, which is a supermartingale under any probability measure $Q \in \mathcal{M}_e(S) \neq \emptyset$. Then, there exists $\alpha \in L(S)$ and C an adapted process, nondecreasing, starting from zero $C_0 = 0$, such that we have the following decomposition:

$$X = X_0 + \int \alpha dS - C \tag{2.14}$$

Remark 2.4.2 *It is important to remark that in the classical Doob-Meyer decomposition theorem of supermartingales X as the difference of a local martingale M and a nondecreasing process C: $X = M - C$, the process C can be chosen to be predictable, and in this case the decomposition is unique. What is remarkable is that the local martingale part $M = X_0 + \int \alpha dS$ can be represented as a stochastic integral with respect to S so that it is a local martingale under any equivalent martingale measure Q. In this sense, decomposition (2.14) is universal. The price to pay is that the increasing process C is in general not predictable as in the Doob-Meyer decomposition but only optional. The processes C have the economic interpretation of cumulative consumption.*

The above decomposition (2.14) implies, as discussed in the rest of this chapter, that the wealth dynamics of the minimal super-hedging portfolio for a contingent claim X_T is given by

$$J_t = \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e(S)} E^Q[X_T | \mathcal{F}_t], \quad 0 \leq t \leq T, \tag{2.15}$$

An analogue result holds for American contingent claims too (see [KQ95, Kra96] for details).

2.4.1 Characterization of the essential supremum of all the possible prices

Let a contingent claim $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. In our financial market, for simplicity, we will always consider the price process of the risk-free asset to be equal to 1 at each date. We make the following assumption

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < \infty \,. \tag{2.16}$$

Let J_t be the essential supremum of the possible prices for X_T at time $t \in [0, T]$ defined by

$$J_t = \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e(S)} E^Q[X_T | \mathcal{F}_t], \quad 0 \leq t \leq T, \tag{2.17}$$

Let us consider the family of adapted processes $\{\Gamma_t^Q \,:\, 0 \leq t \leq T, Q \in \mathcal{M}_e(S)\}$, where Γ_t^Q is defined by

$$\Gamma_t^Q := E^Q[X_T | \mathcal{F}_t], \quad 0 \leq t \leq T, \quad Q \in \mathcal{M}_e(S).$$

This family is well defined thanks to (2.16).

Proposition 2.4.3 *For all $t \in [0,T]$, the set $\{\Gamma_t^Q : Q \in \mathcal{M}_e(S)\}$ is stable with respect to taking finite supremums (and infimums), i.e. for all Q_1, $Q_2 \in \mathcal{M}_e(S)$, there exists $Q \in \mathcal{M}_e(S)$ satisfying* $\max(\Gamma_t^{Q_1}, \Gamma_t^{Q_2}) = \Gamma_t^Q$.

By this property, it follows that for each t, there exists a sequence $Q_p \in \mathcal{M}_e(S)$ so that, almost surely, $\Gamma_t^{Q_p}$ is an increasing sequence of random variables that converges to J_t, that is

$$J_t = \lim_{p\to\infty} \uparrow \Gamma_t^{Q_p} = \lim_{p\to\infty} \uparrow E^{Q_p}[X_T|\mathcal{F}_t].$$

This useful property will allow us to invert supremum and expectation (using the monotone convergence theorem).

Proof. Let Q^0 be a fixed element in the set $\mathcal{M}_e(S)$, i.e. $Q^0 \in \mathcal{M}_e(S)$, and let Z^0 be the martingale density process $dQ^0/dP = Z^0$ and let us define now the process

$$Z_s = \begin{cases} Z_s^0, & s \leq t \\ Z_t^0\big(\frac{Z_s^1}{Z_t^1}\mathbf{1}_A + \frac{Z_s^2}{Z_t^2}\mathbf{1}_{\Omega\setminus A}\big), & t < s \leq T, \end{cases} \tag{2.18}$$

where, for $i = 1, 2$, Z^i is the martingale density of process of Q^i and let A be the $\mathcal{F}_t$ measurable event defined by $A = \{\Gamma_t^{Q_1} \geq \Gamma_t^{Q_2}\}$. It is straightforward to show that Z is also a strictly positive P-martingale with $Z_0 = 1$. We then introduce $Q \sim P$ with Radon-Nikodým density $dQ/dP = Z$. We claim that $Q \in \mathcal{M}_e(S)$, i.e. S is a Q local martingale. In order to check this fact, it is equivalent, thanks to Bayes formula, to prove that ZS is a P local martingale. This latter follows from the fact that the processes Z^iS for $i = 1, 2$ are P local martingales. The desired stability property for supremum follows:

$$\begin{aligned}
\Gamma_t^Q &= E^Q[X_T|\mathcal{F}_t] = E[\frac{Z_T}{Z_t}X_T|\mathcal{F}_t] \\
&= E[\frac{Z_T^1}{Z_t^1}X_T\mathbf{1}_A + \frac{Z_T^2}{Z_t^2}X_T\mathbf{1}_{\Omega\setminus A}|\mathcal{F}_t] \\
&= \mathbf{1}_A E^{Q^1}[X_T|\mathcal{F}_t] + \mathbf{1}_{\Omega\setminus A}E^{Q^2}[X_T|\mathcal{F}_t] \\
&= \mathbf{1}_A\Gamma_t^{Q^1} + \mathbf{1}_{\Omega\setminus A}\Gamma_t^{Q^2} = \max(\Gamma_t^{Q^1}, \Gamma_t^{Q^2}).
\end{aligned}$$

As a consequence, for all $t \in [0,T]$, there exists a sequence $(Q_k^t)_{k\geq 1} \in \mathcal{M}_e(S)$ such that we have

$$J_t := \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} \Gamma_t^Q = \lim_{k\longrightarrow\infty} \uparrow \Gamma_t^{Q_k^t}. \tag{2.19}$$

■

Now we state the universal supermartingale property.

Proposition 2.4.4 *For any $Q_0 \in \mathcal{M}_e(S)$, (J_t) is a supermartingale under Q_0 (that is, $(Z_t^0 J_t)$ is a supermartingale under P where Z^0 is the density process of $\frac{dQ_0}{dP}$).*

Proof. Let us choose an arbitrary element $Q_0 \in \mathcal{M}_e(S)$ with martingale density process $dQ_0/dP = Z^0$, and fix u and t such that $0 \leq u < t \leq T$. We denote by $(Z^{k,t})_{k\geq 1}$ the associated sequence of martingale density processes obtained from the sequence of elements $(Q_k^t)_{k\geq 1} \in \mathcal{M}_e(S)$ given in equation (2.19). For all $k = 1, 2, \cdots$, we remark that the process defined by

$$\tilde{Z}_s^{k,t} = \begin{cases} Z_s^0, & s \leq t \\ Z_t^0 \dfrac{Z_s^{k,t}}{Z_t^{k,t}}, & t < s \leq T, \end{cases} \tag{2.20}$$

is a strictly positive P-martingale with initial value $\tilde{Z}_0^{k,t} = 1$, and let $\tilde{Q}_k^t \sim P$ be the associated probability measure. Moreover, $\tilde{Z}^{k,t} S$ is a P local martingale, and therefore we have $\tilde{Q}_k^t \in \mathcal{M}_e(S)$. For any $k = 1, 2, \cdots$, we have

$$\begin{aligned} E^{Q_0}[\Gamma_t^{Q_k^t} | \mathcal{F}_u] &= E[\frac{Z_t^0}{Z_u^0} \Gamma_t^{Q_k^t} | \mathcal{F}_u] \\ &= E[\frac{Z_t^0}{Z_u^0} E[\frac{Z_T^{k,t}}{Z_t^{k,t}} X_T | \mathcal{F}_t] | \mathcal{F}_u] \\ &= E[\frac{Z_t^0}{Z_u^0} \frac{Z_T^{k,t}}{Z_t^{k,t}} X_T \mid \mathcal{F}_u] \\ &= E[\frac{\tilde{Z}_T^{k,t}}{\tilde{Z}_u^{k,t}} X_T | \mathcal{F}_u] \\ &= E^{\tilde{Q}_k^t}[X_T | \mathcal{F}_u] = \Gamma_u^{\tilde{Q}_k^t}. \end{aligned} \tag{2.21}$$

Equation (2.19) together with the monotone convergence theorem yield

$$\begin{aligned} E^{Q_0}[J_t \mid \mathcal{F}_u] &= \lim_{k \longrightarrow \infty} \uparrow E^{Q_0}[\Gamma_t^{Q_k} | \mathcal{F}_u] = \lim_{k \longrightarrow \infty} \uparrow \tilde{\Gamma}_t^{Q_k^t} \\ &\leq \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e(S)} \Gamma_u^Q = J_u, \end{aligned} \tag{2.22}$$

which proves that J is a Q_0-supermartingale. Setting, $u = 0$, we obtain, thanks to (2.16),

$$E^{Q_0}[J_t] \leq J_0 = \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < \infty \,.$$

which shows that J_t is finite for all $t \in [0, T]$ almost surely. ∎

Proposition 2.4.5 *(J_t) is the smallest supermartingale under Q, for any $Q \in \mathcal{M}_e(S)$, which is equal to X_T at time T (unique up to a null set).*

Proof. Let (J'_t) be a supermartingale under Q, for any $Q \in \mathcal{M}_e(S)$, which is equal to X_T at time T. Then,

$$\forall t \in [0,T] \quad \text{and} \quad Q \in \mathcal{M}_e(S)\,, \quad J'_t \geq E^Q[X_T|\mathcal{F}_t]\,,\; P\,a.s.$$

Hence,

$$\forall t \in [0,T], \quad P\,a.s.\,, \quad J'_t \geq J_t\,.$$

■

We have also the following property

Proposition 2.4.6 *Let $Z^{\nu^\star}$ be the martingale density of $Q^\star \in \mathcal{M}_e(S)$. The following properties are equivalent:*

(i) *$Z^{\nu^\star}$ is optimal, i.e. $\forall t \in [0,T]$, $J_t = E^P[Z_T^{\nu^\star} X_T|\mathcal{F}_t] = E^{Q^\star}[X_T|\mathcal{F}_t]\, P\,a.s.$*

(ii) *$\{Z_t^{\nu^\star} J_t,\, 0 \leq t \leq T\}$ is P-martingale (this is equivalent to $\{J_t,\, 0 \leq t \leq T\}$ is $Q^\star$-martingale).*

Proof. The proof is a direct consequence of the definition of martingale and the fact that $J_T = X_T$. ■

Proposition 2.4.7 *There exists a càd-làg supermartingale still denoted by J_t so that for each $t \in [0,T]$,*

$$J_t = \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e(S)} E^Q[X_T|\mathcal{F}_t], \quad 0 \leq t \leq T,$$

Proof. Let $\mathbb{D} = [0,T] \cap \mathbb{Q}$ where $\mathbb{Q}$ is the set of rational numbers. Because (J_t) is a supermartingale, we have that for almost every $\omega \in \Omega$, the mapping $t \to J_t(\omega)$ defined on $\mathbb{D}$ has at each point t of $[0,T[$ a finite right limit:

$$J_{t+}(\omega) = \lim_{s \in \mathbb{D}, s \downarrow t} J_s(\omega)$$

and at each point of $]0,T]$ a finite left limit ;

$$J_{t-}(\omega) = \lim_{s \in \mathbb{D}, s \uparrow t} J_s(\omega)$$

(see Karatzas and Shreve (1991), Proposition 1.3.14 or Dellacherie and Meyer (1980), Chapter 6).

Note that it is possible to define $J_{t+}(\omega)$ for each $(t,\omega) \in [0,T] \times \Omega$ by:

$$\begin{cases} J_{t+}(\omega) & := & \liminf_{s \in \mathbb{D}, s \downarrow t} J_s(\omega), \quad 0 \le t < T \\ J_{T+}(\omega) & := & J_T(\omega) \end{cases}$$

We show that (J_{t+}) is a càd-làg $\mathcal{F}_{t+}$-supermartingale, which will prove the proposition. We know from Theorem C.0.6 that càd-làg property is equivalent to showing that the function $t \to E^{Q^0}[J_t]$ is right-continuous. Therefore we will prove the right-continuous property of $t \to E^{Q^0}[J_t]$. From (2.22) and taking $u = 0$ we have:

$$E^{Q^0}[J_t] = \lim_{k \to \infty} \uparrow E^{\tilde{Q}^t_k}[X_T] \qquad \forall\, t \in [0,T]. \tag{2.23}$$

Fix $t \in [0,T]$ and let $(t_n)_{n \ge 1} \in [0,T]$ be a decreasing sequence converging to t. From the Q^0-supermatingale property of J, we have

$$\liminf_{n \to \infty} E^{Q^0}[J_{t_n}] \le E^{Q^0}[J_t], \quad \text{and} \quad \limsup_{n \to \infty} E^{Q^0}[J_{t_n}] \le E^{Q^0}[J_t] \tag{2.24}$$

On the other hand, for all $\varepsilon > 0$ there exist by (2.23), $\hat{k} = \hat{k}(\varepsilon) \ge 1$ such that

$$E^{Q^0}[J_t] \le E^{\tilde{Q}^t_{\hat{k}}}[X_T] + \varepsilon. \tag{2.25}$$

Notice that $\tilde{Z}^{\hat{k},t_n}_T$, the Radon-Nikodym density of $\tilde{Q}^{t_n}_{\hat{k}}$ converge a.s. to $\tilde{Z}^{\hat{k},t}_T$, the Radon-Nikodým of the density $\tilde{Q}^t_{\hat{k}}$, as n tends to infinity. Moreover, we have

$$E^{Q^0}[J_{t_n}] \le E^{\tilde{Q}^{t_n}_{\hat{k}}}[X_T] + \varepsilon \le E^{Q^0}[J_{t_n}] + \varepsilon \tag{2.26}$$

where the second inequality follows by (2.23). By Fatou Lemma's, we deduce with 2.26

$$\begin{aligned} E^{Q^0}[J_t] \le E^{Q^0}[J_{t+}] = E^{Q^0}[\liminf_n J_{t_n}] &\le \liminf_n E^{Q^0}[X_{t_n}] && (2.27) \\ &\le \liminf_n E^{\tilde{Q}^{t_n}_{\hat{k}}}[X_T] + \varepsilon && (2.28) \\ &\le \liminf_n E^{Q^0}[J_{t_n}] + \varepsilon && (2.29) \end{aligned}$$

where the second inequality follows by (2.23). Similarly,

$$\begin{aligned} E^{Q^0}[J_t] \le E^{Q^0}[J_{t+}] = E^{Q^0}[\liminf_n J_{t_n}] &\le \liminf_n E^{Q^0}[J_{t_n}] \\ &\le \limsup_n E^{Q^0}[J_{t_n}] \\ &\le \limsup_n E^{\tilde{Q}^{t_n}_{\hat{k}}}[X_T] + \varepsilon \\ &\le \limsup_n E^{Q^0}[J_{t_n}] + \varepsilon \end{aligned} \tag{2.30}$$

Before we conclude, we need to justify the following inequality $E^{Q^0}[J_t] \leq E^{Q^0}[J_{t+}]$. For this, one can show that (J_{t+}) is an $\mathcal{F}_{t+}$-supermartingale under an arbitrary Q_0. Because the filtration is right continuous, (J_{t+}) is an $\mathcal{F}_t$-supermartingale under an arbitrary Q_0. Hence, by Proposition 2.4.5, we have $J_{t+} \geq J_t$. Since ε being arbitrary, this implies $\lim\limits_{n\longrightarrow\infty} E^{Q^0}[J_{t_n}] = E^{Q^0}[J_t]$, i.e. the right continuity of $(E^{Q^0}[J_t])_{t\in[0,T]}$. ■

We have the following property

Proposition 2.4.8 *The process $J = (J_t)_{0\leq t\leq T}$ satisfies the dynamic programming equation*

$$J_s = \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[J_t|\mathcal{F}_s], \quad 0 \leq s \leq t \leq T\,. \tag{2.31}$$

Proof. For $0 \leq s \leq t \leq T$, we have

$$\begin{aligned} J_s &= \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[X_T|\mathcal{F}_s], \\ &= \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[E^Q[X_T|\mathcal{F}_t]|\mathcal{F}_s], \\ &\leq \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[J_t|\mathcal{F}_s] \end{aligned}$$

since

$$J_t = \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[X_T|\mathcal{F}_t], \quad 0 \leq t \leq T.$$

To prove the reverse inequality, and thus (2.31), it suffices to fix an arbitrary $Q \in \mathcal{M}_e(S)$ and show that

$$J_s \geq E^Q[J_t|\mathcal{F}_s]$$

almost surely. This is nothing but the supermartingale property for J, already proved in Proposition (2.4.4). ■

2.4.2 Dual Representation of the Superreplication Cost

In this section we want to see how the optional decomposition Theorem 2.4.1 provides a dual representation of the superreplication problem in 2.5 of a contingent claim $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. Let us consider the càd-làg modification of the process

$$J_t = \operatorname*{ess\,sup}_{Q\in\mathcal{M}_e(S)} E^Q[X_T|\mathcal{F}_t], \quad 0 \leq t \leq T, \tag{2.32}$$

We have shown that $\{J_t\}_{0\leq t\leq T}$ is a supermartingale under any $Q \in \mathcal{M}_e(S)$, therefore we can apply the optional decomposition Theorem 2.4.1 to $\{J_t\}_{0\leq t\leq T}$. The following theorem give the dual representation of the superreplication cost.

Theorem 2.4.9 *Let $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. Then its superreplication cost is equal to*

$$v_0 = J_0 := \sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T]. \tag{2.33}$$

Furthermore if $J_0 < \infty$ i.e. v_0 is finite, then v_0 attains its infimum in (2.5) with a superreplication portfolio strategy $\alpha^\star$ given by the optional decomposition provided in (2.14) of the process J defined in (2.32). That is, there exists a portfolio strategy $\alpha^\star \in L(S)$ such that

$$J_0 + \int_0^T \alpha^\star_s \, dS_s \geq X_T \quad a.s..$$

Moreover, in this case, for any $Q^\star \in \mathcal{M}_e(S)$, the following conditions are equivalent:

(i) *$Q^\star$ achieves the supremum in (2.33).*

(ii) *X_T is attainable: there exists $\alpha \in L(S)$ such that the portfolio $(X_t^{J_0,\alpha} := J_0 + \int_0^t \alpha_s \, dS_s)_{0\leq t\leq T}$ satisfies $X_T^{J_0,\alpha} = J_0 + \int_0^T \alpha_s \, dS_s = X_T$, and the process $\{Z_t^{\nu^\star} X_t^{J_0,\alpha}, 0 \leq t \leq T\}$ is P-martingale (this is equivalent to $\{X_t^{J_0,\alpha}, 0 \leq t \leq T\}$ is $Q^\star$-martingale), where $Z^{\nu^\star}$ is the martingale density of $Q^\star$.*

Proof. We always have that $J_0 := \sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T] \leq v_0$. Indeed, for any $\alpha \in \mathcal{A}(S)$ and any $Q \in \mathcal{M}_e(S)$, the stochastic integral $\int \alpha dS$ is a lower-bounded $Q \in \mathcal{M}_e(S)$ local martingale, and therefore a Q-supermartingale. Hence for any $x \in \mathbb{R}$ with the property that $x + \int_0^T \alpha_t dS_t \geq X_T$ a.s., we obtain

$$E^Q[X_T] \leq E^Q[x + \int_0^T \alpha_t dS_t] \leq x \,,$$

since $E^Q[\int_0^T \alpha_t dS_t] \leq 0$, because of the supermartingale property. Thus, $E^Q[X_T] \leq x$ for all $Q \in \mathcal{M}_e(S)$. We conclude from the definition of v_0 that we have the following inequality

$$J_0 = \sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T] \leq v_0. \tag{2.34}$$

Conversely, we want to prove that $J_0 = \sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T] \geq v_0$ which is the delicate part of the proof and requires the use of the optional decomposition theorem. If

$$\sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T] = \infty \,,$$

there is nothing to prove. So let us now assume that

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < \infty \,.$$

This makes it possible to apply the optional decomposition theorem to the càd-làg modification, still denoted by J, and obtain the existence of a process $\alpha^\star \in L(S)$, and an adapted nondecreasing process C, with $C_0 = 0$ such that

$$J_t = J_0 + \int_0^t \alpha_s^\star dS_s - C_t \quad 0 \leq t \leq T \,, \qquad \text{a.s.} \tag{2.35}$$

Since J and C are nonnegative, this last relation shows that $\int \alpha^\star dS$ is lower-bounded (by $-J_0$), and so $\alpha^\star \in \mathcal{A}(S)$. Furthermore, the relation (2.35) for $t = T$ implies

$$J_T = X_T \leq J_0 + \int_0^T \alpha_s^\star dS_s \,, \quad \text{a.s.} \tag{2.36}$$

This shows by definition of v_0 that

$$v_0 \leq J_0 = \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T]. \tag{2.37}$$

This concludes the proof of the first part of the theorem. It remains now to show the second part of the theorem. Let (iii) be the condition that the process $\{Z_t^{\nu^\star} J_t,\, 0 \leq t \leq T\}$ is a P-martingale.

We show that conditions (i), (ii) and (iii) are equivalent. The P-supermartingale $Z^{\nu^\star} J$ is a P-martingale if and only if $J_0 = E[Z_T^{\nu^\star} J_T] \Longleftrightarrow J_0 = E[Z_T^{\nu^\star} X_T] \Longleftrightarrow (i)$.

On the other hand, (iii) implies that from 2.35 we have $C \equiv 0$, hence $X_t^{J_0,\alpha} = J_t = J_0 + \int_0^t \alpha^\star \, dS_s$. Thus, (ii) is satisfied with $\alpha = \alpha^\star$. On the other hand, suppose that (ii) holds. Then, $J_0 = E[Z_T^{\nu^\star} X_T]$ and (i) holds. ■

Thanks to the dual representation (2.33) of the superreplication cost, we shall obtain in the next Section a very useful characterization of the sets $\mathcal{C}(x)$ as stated in Corollary 2.5.1.

2.5 Dual Space Characterisation

The following Corollary provides a representation and characterisation of $\mathcal{C}(x)$ in terms of some dual space of probability measures.

Corollary 2.5.1 *For all $x \in \mathbb{R}_+$, we have*

$$\mathcal{C}(x) = \{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) : \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq x\}. \tag{2.38}$$

Consequently, $\mathcal{C}(x)$ is convex, solid and closed for the topology of the convergence in measure i.e. if $(X^n)_{n\geq 1}$ is a sequence in $\mathcal{C}(x)$ converging a.s. to $\hat{X}_T$, then $\hat{X}_T \in \mathcal{C}(x)$. Moreover, $\mathcal{C}(x)$ is a bounded subset of $L^0_+(\Omega, \mathcal{F}_T, P)$ and contains the constant random variable $X_T = x$.*

Proof. Let $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$. For any admissible strategy $\alpha \in \mathcal{A}(S)$, the corresponding value process $\int \alpha dS$ is a supermartingale for any $Q \in \mathcal{M}_e$. Hence, if α satisfies

$$x + \int_0^T \alpha_t dS_t \geq X_T \quad a.s.\,,$$

we immediately have

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq x\,.$$

This means that

$$\mathcal{C}(x) \subseteq \{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) : \sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq x\}.$$

For the reverse implication, which is the more difficult part of the proof, consider any $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$ such that

$$\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] \leq x\,.$$

Denote by J a right-continuous version of the process

$$J_t = \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e(S)} E^Q[X_T | \mathcal{F}_t], \quad 0 \leq t \leq T\,.$$

By hypothesis,

$$J_0 \leq x. \tag{2.39}$$

(J_t) is a supermartingale under any pricing rule $Q \in \mathcal{M}_e$. On account of the optional decomposition theorem, this yields the existence of $\alpha \in L(S)$ (which is also in $\mathcal{A}(S)$ since $\int \alpha dS \geq -J_0$) and an increasing optional process C satisfying $C_0 = 0$ such that for all $t \in [0, T]$

$$J_t = J_0 + \int_0^t \alpha_s dS_s - C_t, \quad a.s.$$

*A subset $\mathcal{C} \in L^0_+(\Omega, \mathcal{F}_T, P)$ is said to be solid, if $0 \leq h \leq f$ and $f \in \mathcal{C}$ implies that $h \in \mathcal{C}$.

holds. Therefore, we can estimate X_T almost surely by

$$\begin{aligned} X_T &= J_T = J_0 + \int_0^T \alpha_s dS_s - C_T \quad (2.40)\\ &\leq J_0 + \int_0^T \alpha_s dS_s \quad (2.41)\\ &\leq x + \int_0^T \alpha_s dS_s\,, \quad (2.42) \end{aligned}$$

by taking (2.39) into account. We conclude, this time

$$\{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) : \sup_{Q\in\mathcal{M}_e(S)} E^Q[X_T] \leq x\} \subseteq \mathcal{C}(x)\,.$$

The convexity and solidity of $\mathcal{C}(x)$ are rather obvious. Let $(X^n)_{n\geq 1}$ be a sequence in $\mathcal{C}(x)$ converging to $\hat{X}_T$ a.s.. Take $Q \in \mathcal{M}_e(S)$. By Fatou's lemma, we have

$$E^Q[\hat{X}_T] \leq \liminf_{n\to\infty} E^Q[X^n] \leq x$$

and so $\hat{X}_T \in \mathcal{C}(x)$. This establish the closedness property of $\mathcal{C}(x)$. ∎

Remark 2.5.2 *It is easy to see that the following useful properties hold:*

1. *For* $0 < x_1 < x_2$*, we have* $\mathcal{C}(x_1) \subseteq \mathcal{C}(x_2)$*.*
2. *For* $x_1, x_2 \in \mathbb{R}^+$ *and* $\varepsilon \in (0,1)$*, we have*

$$\varepsilon\, X_T^{x_1} + (1-\varepsilon)\, X_T^{x_2} \in \mathcal{C}(\varepsilon\, x_1 + (1-\varepsilon)\, x_2)$$

where $X_T^{x_i} \in \mathcal{C}(x_i)$*,* $i = 1, 2$*.*

Corollary 2.5.1 provides an extremely useful and simple characterization of the set $\mathcal{C}(x)$: in order to know if a European contingent claim can be dominated almost surely from an initial capital x and an admissible portfolio strategy, it is *necessary and sufficient* to test if its expectation under any martingale probability measure is less or equal to x. Mathematically, as we will see in the next Chapter, this characterization is the crucial step in transforming a dynamic expected utility optimization problem into a static one and it forms the starting point for the duality methods approach in the resolution of the utility maximization problem from terminal wealth. Moreover, thanks to this dual space characterization, the closure property of the set $\mathcal{C}(x)$ in $L^0_+(\Omega, \mathcal{F}_T, P)$ is easily obtained, which was not evident from its original (primal) definition (2.11).

2.6 Superhedging Theorem of European Options

Let $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$ be a European contingent claim. The fundamental result in the literature on super-hedging is the dual characterization of the set $\mathcal{D}^{X_T}$ of *hedging endowment*

$$D^{X_T} := \left\{ x \in \mathbb{R} : \exists \alpha \in \mathcal{A}(S),\ \ x + \int_0^T \alpha_t dS_t \geq X_T \text{ a.s.} \right\}. \tag{2.43}$$

i.e. D^{X_T} is the set of capitals starting or initial endowments $x \in \mathbb{R}$ from which one can super-replicate the pay-off of an ECC X_T with maturity T by the terminal value of a self-financing and admissible portfolio.

Theorem 2.6.1 ***(Superhedging Theorem of European Options)***
Suppose that $E^Q[X_T] < \infty$ for every $Q \in \mathcal{M}_e(S)$. Then

$$D^{X_T} = \left\{ x \in \mathbb{R} \,:\, x \geq E^Q[X_T]\,, \forall Q \in \mathcal{M}_e(S) \right\} =: [\bar{x}, \infty)\,. \tag{2.44}$$

where $\bar{x} = v_0(X_T)$.

Proof. We note that the following inclusion

$$D^{X_T} \subseteq \left\{ x \in \mathbb{R} \,:\, x \geq E^Q[X_T]\,, \forall Q \in \mathcal{M}_e(S) \right\}. \tag{2.45}$$

is obvious: if $x + \int_0^T \alpha_t dS_t \geq X_T$ then $x \geq E^Q[X_T]$ for every $Q \in \mathcal{M}_e(S)$. To show the opposite inclusion, one suppose that $\sup_{Q \in \mathcal{M}_e(S)} E^Q[X_T] < \infty$ (otherwise both sets are empty). We then apply the optional decomposition theorem as in the proof of Corollary 2.5.1. ∎

2.7 Itô processes and Brownian filtration framework

Remember that in our financial market, for simplicity, we will always consider the price process of the risk-free asset to be constant and equal to 1 at each date. The asset price process $S = (S^1, S^2, \cdots, S^n)$ model will follows the dynamics described by:

$$dS_t = \mu_t dt + \sigma_t dW_t \tag{2.46}$$

where W is a d-dimensional standard Brownian motion on the complete filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ the natural filtration of W, and $d \geq n$

(i.e. the number of risk factors d is larger than the number of stocks n), the drift μ is a n-dimensional progressively measurable process, the volatility σ is a progressively measurable $(n \times d)$ matrix-valued process and we have the following integrability condition $\int_0^T |\mu_t|dt + \int_0^T |\sigma_t|^2 dt < \infty \, a.s.$ For all $t \in [0,T]$, the $(n \times d)$ matrix-valued σ_t is assumed of full rank equal to n. The square $n \times n$ matrix-valued covariance process $\sigma_t \sigma_t'$, is thus invertible, and we finally define the risk-premium process λ, a d-dimensional progressively measurable process, by:

$$\lambda_t := \sigma_t'(\sigma_t \sigma_t')^{-1} \mu_t, \qquad 0 \le t \le T. \tag{2.47}$$

From (2.47), we have $\sigma_t \lambda_t = \mu_t$. For simplicity, the process λ will be supposed (see Remark 3.4.13) to be bounded.

In the current setting of a Brownian filtration, the equivalent martingale measures $Q \in \mathcal{M}_e(S)$ can be parametrized quite explicitly. Girsanov theorem may be used to remove the drift of S and obtain an equivalent martingale measure. For example, since $(\lambda_t)_{0 \le t \le T}$ is bounded,

$$\frac{d\widehat{Q}}{dP} := \mathcal{E}\Big(- \int \lambda \cdot dW \Big)_T. \tag{2.48}$$

clearly defines a EMM $\widehat{Q}$, which is known as the minimal martingale measure, see Remark 2.7.4. Under $\widehat{Q}$, the process

$$\widehat{W}_0 = 0\,, \qquad d\widehat{W}_t = dW_t + \lambda_t dt\,, \quad 0 \le t \le T. \tag{2.49}$$

is a standard Brownian motion. Moreover, (2.46) becomes $dS_t = \sigma_t d\widehat{W}_t$.

Remark 2.7.1 *In order to insure a positive price process S, we traditionally take an Itô dynamics for S in the form:*

$$dS_t = diag(S_t)(\tilde{\mu}_t dt + \tilde{\sigma}_t dW_t), \qquad 0 \le t \le T. \tag{2.50}$$

where $diag(S_t)$ denotes the diagonal $n \times n$ matrix with diagonal elements S_t^i. The Black-Scholes model and stochastic volatility models are particular examples of (2.50). Notice that the model (2.50) is a special case of the model (2.46) with

$$\mu_t = diag(S_t)\tilde{\mu}_t \qquad \textit{and} \qquad \sigma_t = diag(S_t)\tilde{\sigma}_t\,. \tag{2.51}$$

We aim now to establish an explicit description of the set of equivalent (local) martingale measures $\mathcal{M}_e(S)$ under the above framework which will be extremely useful in the remaining parts of the thesis. For this, let us introduce the set

$$K(\sigma) = \{\nu \in L^2_{loc}(W) : \sigma\nu = 0, \;\; [0,T] \times \Omega, \;\; dt \otimes dP \; a.e.\}. \tag{2.52}$$

For any $\nu \in K(\sigma)$ we define for $0 \le t \le T$, the exponential local martingale

$$Z^\nu_t := \mathcal{E}\Big(- \int (\lambda + \nu) \cdot dW \Big)_t = \exp(-\int_0^t (\lambda_s + \nu_s) \cdot dW_s - \frac{1}{2} \int_0^t |\lambda_s + \nu_s|^2 ds) \,. \tag{2.53}$$

Notice that the facts that the $n \times n$ matrix-valued covariance process $\sigma\sigma'$ is invertible, $\sigma\lambda = \mu$ and $\sigma\nu = 0$, imply that $\lambda'\nu = 0$, in other words that λ and ν are orthogonal, therefore we have $|\lambda + \nu|^2 = |\lambda|^2 + |\nu|^2$ and Equation (2.53) becomes

$$Z^\nu_t = \exp(-\int_0^t (\lambda_s + \nu_s) \cdot dW_s - \frac{1}{2} \int_0^t (|\lambda_s|^2 + |\nu_s|^2) ds) \,. \tag{2.54}$$

We also define the subset $K_m(\sigma)$ of $K(\sigma)$ by

$$K_m(\sigma) = \{\nu \in K(\sigma) : Z^\nu \text{ is a true martingale}\} \,. \tag{2.55}$$

Remark 2.7.2 *It is well-known that $E[Z^\nu_T] = 1$ is a necessary and sufficient condition ensuring that Z^ν is a true martingale. However, a sufficient condition for Z^ν to be a true martingale is given by the Novikov criterion:*

$$E[\exp(\frac{1}{2} \int_0^T (|\lambda_s|^2 + |\nu_s|^2) \, ds)] < \infty \,. \tag{2.56}$$

Since we assumed λ to be bounded, the Novikov condition (2.56) holds for any bounded process ν. In particular, the null process $\nu = 0$ belongs to $K_m(\sigma)$.

For any element $\nu \in K_m(\sigma)$, one can define an equivalent probability measure $P^\nu \sim P$ with martingale density process $dP^\nu/P = Z^\nu$ such that the process

$$W^\nu = W + \int (\lambda + \nu) dt \tag{2.57}$$

is a standard P^ν-Brownian motion thanks to Girsanov theorem.

The following proposition due to El Karoui and Quenez [KQ95] give us the desired identification of the set of EMM $\mathcal{M}_e(S)$ via exponential densities.

Proposition 2.7.3 *We have the following*

$$\mathcal{M}_e(S) = \{P^\nu : \nu \in K_m(\sigma)\} \tag{2.58}$$

Proof. (i) We first prove that $\{P^\nu : \nu \in K_m(\sigma)\} \subset \mathcal{M}_e(S)$. Since by construction, $\lambda\sigma = \mu$ and for all $\nu \in K_m(\sigma)$, we have $\sigma\nu = 0$ it follows that the dynamics of S under P^ν is given by

$$dS_t = \sigma_t[\lambda_t dt + dW_t] = \sigma_t[\lambda_t dt + dW_t^\nu - (\lambda_t + \nu_t)dt] = \sigma_t dW_t^\nu. \tag{2.59}$$

Thus, S is a P^νlocal martingale, i.e. $P^\nu \in \mathcal{M}_e(S)$.

(ii) Conversely, let $Q \in \mathcal{M}_e(S)$, in particular $Q \sim P$ and therefore let Z denotes its strictly positive martingale density process. By the martingale Itô representation theorem there exist $\rho \in L^2_{loc}(W)$ such that

$$Z_t = \mathcal{E}\Big(-\int \rho \cdot dW\Big)_t = \exp[-\int_0^t \rho_s \cdot dW_s - \frac{1}{2}\int_0^t |\rho_s|^2 ds], \qquad 0 \le t \le T\,. \tag{2.60}$$

The process $B^\rho = W + \int \rho dt$ is a Q-Brownian motion thanks to Girsanov theorem. Therefore, the process S has the following dynamics under Q:

$$dS_t = (\mu_t - \sigma_t\rho_t)dt + \sigma_t dB_t^\rho, \qquad 0 \le t \le T. \tag{2.61}$$

Since S is a Q-local martingale , the drift should be identically zero, hence

$$\mu = \sigma\rho \quad \text{on} \quad [0,T] \times \Omega, \;\; dt \otimes dP\; a.e. \tag{2.62}$$

By defining $\nu := \rho - \lambda$ and using the fact that $\sigma\lambda = \mu$ it follows that $\sigma\nu = 0$ and therefore $\nu \in K(\sigma)$. To conclude, we notice that $Z^\nu = Z$ (which is a true martingale), so $\nu \in K_m(\sigma)$ and we finally have the desired result $Q = P^\nu$. ■

Remark 2.7.4 *1. The inclusion $\{P^\nu : \nu \in K_m(\sigma)\} \subseteq \mathcal{M}_e(S)$ remains valid without the Brownian filtration hypothesis, as seen from the Part (i) of the above proof.*

2. In fact, the Novikov condition (2.56) holds for any bounded process ν and when λ is not necessarily assumed to be bounded but satisfies the Novikov criterion $E[\exp(\frac{1}{2}\int_0^T |\lambda_u|^2 du)] < \infty$. In particular, the null process $\nu = 0$ is an element of the set $K_m(\sigma)$. The associated EMM P^0 characterized in term of risk premia process λ is called the Föllmer-Schweizer minimal martingale measure.

3. *The above observation also indicates that once λ satisfies the Novikov condition, $P^0 \in \mathcal{M}_e(S)$ and therefore $\mathcal{M}_e(S) \neq \emptyset$, in other words the market is arbitrage free. In the case where Z^0 is not a true martingale, i.e. $P^0 \notin \mathcal{M}_e(S)$, we do not necessarily have the no-arbitrage condition $\mathcal{M}_e(S) \neq \emptyset$, and $\mathcal{M}_e(S) \neq \emptyset$ is equivalent to the existence of at least an element ν in $K_m(\sigma)$.*

We end this Chapter by a proof of the optional decomposition theorem in the setting 2.7 of Itô processes and Brownian filtration. In fact, in this case, the process C in the decomposition is predictable.

Theorem 2.7.5 *Let $X \geq 0$ be a càd-làg process which is a supermartingale under any martingale measure P^ν, $\nu \in K_m(\sigma)$. Then, X admits a decomposition under the form*

$$X = X_0 + \int \alpha dS - C \tag{2.63}$$

where $\alpha \in L(S)$ and C is a nondecreasing predictable process, $C_0 = 0$.

Proof. The application of the classical Doob-Meyer decomposition theorem to the non-negative super-martingale X under P^ν for $\nu \in K_m(\sigma)$ yields the following decomposition

$$X_t = X_0 + M_t^\nu - A_t^\nu, \qquad 0 \leq t \leq T \tag{2.64}$$

where M^ν is a local martingale under P^ν, starting from zero $M_0^\nu = 0$ and A^ν is a predictable nondecreasing process, integrable (under P^ν), with $A_0^\nu = 0$. By the (local) martingale Itô representation theorem under P^ν, there exist a $\psi^\nu \in L^2_{loc}(W^\nu)$ such that

$$X_t = X_0 + \int_0^t \psi_u^\nu dW_u^\nu - A_t^\nu, \qquad 0 \leq t \leq T \tag{2.65}$$

Choose some element in $\mathcal{M}_e(S)$, say P^0 for convenience, and compare the decompositions (2.65) of X under P^ν, and P^0. Note that $W^\nu = W^0 + \int \nu dt$ and by identifying the (local) martingale and predictable finite variation parts, we get a.s.

$$\begin{aligned} \psi_t^\nu &= \psi_0^\nu, \qquad 0 \leq t \leq T, & (2.66) \\ A_t^\nu - \int_0^t \nu_u' \psi_u^\nu du &= A_t^0, \qquad 0 \leq t \leq T, & (2.67) \end{aligned}$$

for all $\nu \in K_m(\sigma)$.

We define the $\mathbb{R}^n$-valued progressively measurable process α by

$$\alpha_t = (\sigma_t \sigma_t')^{-1} \sigma_t \psi_t^0, \qquad 0 \leq t \leq T. \tag{2.68}$$

Observe that $\int_0^T |\alpha_t' \mu_t| dt = \int_0^T |\lambda_t' \psi_t^0| dt < \infty$ and $\int_0^T |\alpha_t' \sigma_t|^2 dt = \int_0^T |\psi_t^0|^2 dt < \infty$ a.s. and so $\alpha \in L(S)$. By writing $\eta_t = \psi_t^0 - \sigma_t' \alpha_t$ we have $\int_0^T |\eta_t|^2 dt < \infty$ a.s., $\sigma\eta = 0$, and so $\eta \in K(\sigma)$. In fact, we obtained the decomposition of ψ^0 on $\mathrm{Im}(\sigma')$ and its orthogonal space $K(\sigma)$.

$$\psi_t^0 = \sigma_t' \alpha_t + \eta_t, \qquad 0 \leq t \leq T. \tag{2.69}$$

We now want to prove that

$$\eta = 0\,, \text{ on } \;[0,T] \times \Omega,\;\; dt \otimes dP \; a.e. \tag{2.70}$$

Let us consider, for any $n \in \mathbb{N}$ the process

$$\nu_t^n = -n \frac{\eta_t}{|\eta_t|} \mathbf{1}_{\eta_t \neq 0}, \qquad 0 \leq t \leq T. \tag{2.71}$$

Then ν^n is bounded, and lies in $K_m(\sigma)$. From (2.66) for ν^n and using also (2.68) we get,

$$A_T^{\nu^n} = A_T^0 - n \int_0^T |\eta_t| \mathbf{1}_{\eta_t \neq 0} dt. \tag{2.72}$$

Assuming that (2.70) does not hold, we get for n large enough

$$E^{P_0}[A_T^{\nu^n}] = E^{P_0}[A_T^0] - n E^{P_0}[\int_0^T |\eta_t| \mathbf{1}_{\eta_t \neq 0} dt] < 0. \tag{2.73}$$

since $E^{P_0}[A_T^0] < \infty$, which leads to a contradiction because $E^{P_0}[A_T^{\tilde{\nu}}] \geq 0$. This implies the desired result:

$$\eta = 0 \text{ on } \;[0,T] \times \Omega,\;\; dt \otimes dP \; a.e.$$

By recalling the dynamics 2.59 of S under P^0 , and writing $C = A^0$, the decomposition 2.60 of X under P^0 is written as

$$X = X_0 + \int \alpha' \sigma dW^0 - A^0 = X_0 + \int \alpha dS - C \tag{2.74}$$

and the proof is complete.

∎

Chapter 3

Duality for the Utility Maximisation Problem

This chapter forms the theoretical core of this thesis. We start by giving a rigorous formulation of the portfolio expected utility optimization problem and specifying some required assumptions on the utility function U as well as on the value function $v(\cdot)$. We then state the equivalent static problem and give an existence and uniqueness result for the solution to the static problem. In Section (3.3), we will study and explain the conjugate duality method for utility from terminal wealth only (without consumption). We conclude this chapter, by stating and proving the main theorem due to Kramkov and Schachermayer on the portfolio optimization problem. The detailed proof is divided into a sequence of lemmas and propositions of independent interest.

3.1 Formulation of the Portfolio Expected Utility Optimization Problem

In this Section we formulate the portfolio expected utility maximization problem under the financial market model mentioned in Section 2.7. Remember that we work with Itô processes and under a Brownian filtration framework. The investor, endowed with some initial wealth, trades dynamically (i.e. continuously in time) the asset price process $S = (S^1, S^2, \cdots, S^n)$. Her objective is to maximize the expected utility function which models her individual preferences as well as her attitude towards the market risk. The utility of an investor with wealth x is given by a function $U(x)$. As discussed in Section 1.3 we make the following standard assumptions on the utility function. The function $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ is continuous on its domain $\mathrm{dom}(U) = \{x \in \mathbb{R} : U(x) > -\infty\}$, differentiable, strictly increasing and strictly concave on the interior of its domain. There is no loss of generality in assuming that $U(\infty) > 0$ (if not we may add a constant). We impose the following condition on U:

Assumption 3.1.1 *[Negative wealth not allowed]*

$$int(dom(U)) = (0, \infty)\,, \tag{3.1}$$

which means that the investor will avoid bankruptcy at all cost. She will assign infinitely negative utility to the possibility of zero or negative wealth.

The portfolio expected utility maximization problem from terminal wealth is then formulated as

$$v(x) = \sup_{\alpha \in \mathcal{A}(S)} E[U(x + \int_0^T \alpha_s dS_s)], \qquad x > 0. \tag{3.2}$$

Lastly, in order to rule out trivial cases, the value function (3.2) is supposed to be non-degenerate:

Assumption 3.1.2 *[Non-degeneracy]*

$$v(x) < \infty\,, \qquad \textit{for some } x > 0. \tag{3.3}$$

Remark 3.1.3 **(i)** *from the fact that U is increasing and concave on its domain, Assumption (3.1.2) is equivalent to:*

$$v(x) < \infty\,, \qquad \textit{for all } x > 0. \tag{3.4}$$

(ii) *We may also introduce and study the dynamical value function of problem (3.2) defined as, for $x > 0$ and $0 \leq t \leq T$*

$$v(t, x) = \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}(S)} E[U(x + \int_t^T \alpha_s dS_s)|\mathcal{F}_t]. \tag{3.5}$$

Using Itô-Ventzell formula, Mania et al. [MT03] derived a Backward Stochastic Partial Differential Equation (BSPDE) for the dynamical value function.

3.2 Equivalent Static Problem and General Existence Result

In this section, we begin by formulating the equivalent static problem and use this result to obtain a proof of the existence (and uniqueness) of a solution to the expected utility maximization problem (3.2).

3.2.1 Equivalent Static Problem

First, observe that Assumption 3.1.1 implies that $U(x) = -\infty$ for $x < 0$, it is therefore enough to consider in the supremum of (3.2) the controls $\alpha \in \mathcal{A}(S)$ leading to nonnegative terminal wealth $x + \int_0^T \alpha_s dS_s \geq 0 \, a.s.$. Let us then define the set

$$\mathcal{W}(x) := \{W = (W_t)_{0 \leq t \leq T} \,:\, W_t = x + \int_0^t \alpha_s dS_s \text{ with } \alpha \in \mathcal{A}(S) \text{ and } W_T \geq 0 \; a.s.\}.$$

The following lemma formulates the dynamic problem

$$v(x) = \sup_{W \in \mathcal{W}(x)} E[U(x + \int_0^T \alpha_s dS_s)], \tag{3.6}$$

into an equivalent static one.

Lemma 3.2.1 *We have*

$$v(x) = \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)], \tag{3.7}$$

where the set

$$\mathcal{C}(x) = \left\{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) \,:\, 0 \leq X_T \leq W_T \; a.s. \quad \text{for some} \quad W \in \mathcal{W}(x)\right\},$$

was defined in (2.11).

(1) If $W^\star \in \mathcal{W}(x)$ solves (3.6), then $X_T^\star = W_T^\star$ solves (3.7)

(2) Conversely, if $X_T^\star \in \mathcal{C}(x)$ solves (3.7), then there exists a $W^\star \in \mathcal{W}(x)$, s.t. $X_T^\star \leq W_T^\star$ which solves (3.6).

In other words, we pass from the set of processes $\mathcal{W}(x)$ to the set $\mathcal{C}(x)$ of *random variables* dominated by the final outcomes W_T.

Proof. Let $W \in \mathcal{W}(x)$. Then $W_T \in \mathcal{C}(x)$ and so $E[U(W_T)] \leq \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)]$, hence

$$v(x) \leq \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)]\,. \tag{3.8}$$

Conversely, given $X_T \in \mathcal{C}(x)$, there exists $W \in \mathcal{W}(x)$ such that $W_T \geq X_T$ a.s. Since U is increasing on $(0, \infty)$, we deduce that $E[U(X_T)] \leq E[U(W_T)]$ and so by (3.6)

$$\sup_{X_T \in \mathcal{C}(x)} E[U(X_T)] \leq v(x) \tag{3.9}$$

which proves (3.7).

(1) Suppose that $W^\star \in \mathcal{W}(x)$ solves (3.6). Then $X_T^\star = W_T^\star \in \mathcal{C}(x)$ and we have

$$v(x) = E[U(W_T^\star)] = E[U(X_T^\star)]$$

which shows that $X_T^\star$ solves (3.7).

(2) Suppose that $X_T^\star \in \mathcal{C}(x)$ solves (3.7). Then there exists a $W^\star \in \mathcal{W}(x)$, $s.t.\, X_T^\star \leq W_T^\star$ a.s.. Since U is increasing on $(0, \infty)$, we have

$$v(x) = E[U(X_T^\star)] \leq E[U(W_T^\star)]$$

which shows that $W^\star$ solves (3.6). Moreover, we can add that, U being (strictly) increasing, we necessarily have, in this case $X_T^\star = W_T^\star$. In other words, there exists $\hat{\alpha} \in \mathcal{A}(S)$ such that $X_T^\star = x + \int_0^T \hat{\alpha}_t dS_t$ and $\hat{\alpha}$ is a solution of (3.6). ∎

In the next Subsection we show the existence of a solution to the static problem (3.7) thanks to the dual characterization of $\mathcal{C}(x)$ established in Corollary (2.5.1), and more precisely from its closure property in $L^0_+(\Omega, \mathcal{F}_T, P)$. The idea behind is to consider a maximizing sequence $(X^n)_{n \geq 1}$ for (3.7), and to use Theorem A.3.2 of compactness in $L^0_+(\Omega, \mathcal{F}_T, P)$, which allows us, up to a convex combination, to obtain a limit a.s. $\hat{X}_T$ of X^n, and then to pass to the limit in $E[U(X^n)]$. The sequence $(U^+(X^n))_{n \geq 1}$ needs to be uniform integrable in order to proceed to the limit. We therefore require the following condition on the value function

Assumption 3.2.2

$$\lim_{x \to \infty} \sup \frac{v(x)}{x} \leq 0. \tag{3.10}$$

This assumption may look a priori weird and difficult to verify in practice since it is stated on the unknown value function $v(\cdot)$ to be determined. In fact, we shall see in the proof of Theorem 3.2.3 below that it is exactly the necessary condition to ensure the convergence of the sequence $(E[U(X^n)])_{n\geq 1}$ to $E[U(\hat{X}_T)]$. On the other hand, in Remark 3.3.3 some practical conditions bearing directly on the utility function U, which ensure (3.10) are given.

3.2.2 Existence and Uniqueness

The following theorem gives the existence and uniqueness of the solution to the static problem (3.7):

Theorem 3.2.3 *Let U be a utility function satisfying assumptions (3.1.1), (3.1.2) and (3.2.2). Then, for all $x > 0$, there exists a unique solution $\hat{X}_T^x$ to problem $v(x)$ in (3.7).*

Proof. Let $x > 0$ and $(X^n)_{n\geq 1}$ be a maximising sequence in $\mathcal{C}(x)$ for $v(x) < \infty$. It means that

$$\lim_{n\longrightarrow\infty} E[U(X^n)] = v(x) < \infty. \tag{3.11}$$

From the compactness Theorem A.3.2 in $L^0_+(\Omega, \mathcal{F}_T, P)$, there exists a convex combination $\hat{X}^n \in conv(X^n, X^{n+1}, \cdots)^*$, which is still in the convex set $\mathcal{C}(x)$ and such that $\hat{X}^n$ converge a.s. to some non-negative random variable $\hat{X}_T^x$. Since, $\mathcal{C}(x)$ is closed for the convergence in measure, we have $\hat{X}_T^x \in \mathcal{C}(x)$. By the concavity of U and from 3.11 we also have

$$\lim_{n\longrightarrow\infty} E[U(\hat{X}^n)] = v(x) < \infty. \tag{3.12}$$

We have,

$$\lim_{n\longrightarrow\infty} E[U(X^n)] = v(x) \Longleftrightarrow v(x) - \frac{1}{n} \leq E[U(X^n)] \leq v(x)\ \forall n \geq 1. \tag{3.13}$$

Since U is concave, we have

$$E[U(\hat{X}^n)] = E[U(\sum_{m\geq n} a_{m,n} X^m)] \geq \sum_{m\geq n} a_{m,n} E[U(X^m)], \tag{3.14}$$

$^*\hat{Y}^n \in conv(Y^n, Y^{n+1}, \cdots)$, means that for any $n \geq 1$, $\hat{Y}^n = \sum_{j=0}^{k(n)} a_j^n Y^{n+j}$, with positive constant a_j^n such that $\sum_{j=0}^{k(n)} a_j^n = 1$.

where $a_{m,n} \geq 0$ and $\sum\limits_{m \geq n} a_{m,n} = 1$ and only finitely many $a_{m,n}$ are nonzero, which gives

$$v(x) \geq E[U(X^m)] \geq v(x) - \frac{1}{m} \geq v(x) - \frac{1}{n} \qquad m \geq n. \tag{3.15}$$

which implies

$$v(x) \geq \min_{m \geq n} E[U(X^m)] \geq v(x) - \frac{1}{n}\,, \tag{3.16}$$

and therefore

$$\lim_{n \to \infty} \min_{m \geq n} E[U(X^m)] = v(x), \tag{3.17}$$

We also have

$$E[U(\hat{X}^n)] \geq \sum_{m \geq n} a_m E[U(X^m)] \geq \min_{m \geq n} E[U(X^m)], \tag{3.18}$$

thus, with $\hat{X}^n \in \mathcal{C}(x)$

$$\min_{m \geq n} E[U(X^m)] \leq E[U(\hat{X}^n)] \leq v(x), \tag{3.19}$$

which implies, by taking the limits in both side of (3.19)

$$v(x) \leq \lim_{n \to \infty} E[U(\hat{X}^n)] \leq v(x), \tag{3.20}$$

and we obtain the result

$$\lim_{n \to \infty} E[U(\hat{X}^n)] = v(x) < \infty. \tag{3.21}$$

Let us denote by U^+ and U^- the positive and the negative parts of U and from 3.12 we have $\sup\limits_n E[U^-(\hat{X}^n)] < \infty$ and $\sup\limits_n E[U^+(\hat{X}^n)] < \infty$. Moreover, by Fatou's Lemma, we have

$$\begin{aligned}
\liminf_{n \longrightarrow \infty} E[U^-(\hat{X}^n)] &\geq E[\liminf_{n \longrightarrow \infty} U^-(\hat{X}^n)] \\
&= E[U^-(\liminf_{n \longrightarrow \infty} \hat{X}^n)] \\
&= E[U^-(\hat{X}^x_T)].
\end{aligned}$$

The optimality of $\hat{X}^x_T$, i.e. $v(x) = E[U(\hat{X}^x_T)]$, is thus obtained if and only if we can show that

$$\lim_{n \longrightarrow \infty} E[U^+(\hat{X}^n)] = E[U(\hat{X}^x_T)], \tag{3.22}$$

which, thanks to Theorem A.1.2, is equivalent to the uniformly integrability of the sequence $U^+(\hat{X}^n)_{n\geq 1}$ (notice that if $U(\infty) \leq 0$, then $U^+ \equiv 0$ and this case is settled). Now, remember that in Section 3.1 we supposed that $U(\infty) > 0$, and let us define

$$x_0 := \inf\{x > 0 : U(x) \geq 0\} < \infty.$$

We proceed by contradiction. Assume that the sequence $U^+(\hat{X}^n)_{n\geq 1}$ is not uniformly integrable. Then, there is $\rho > 0$ with the property that

$$\lim_{n\longrightarrow\infty} E[U^+(\hat{X}^n)] = E[U(\hat{X}^x_T)] + 2\rho.$$

From Corollary (A.1.3), and considering the subsequence still denoted by $(\hat{X}^n)_{n\geq 1}$, we can find $\mathcal{F}_T$ measurable partition $(B^n)_{n\geq 1}$ of Ω such that

$$E[U^+(\hat{X}^n)\mathbf{1}_{B^n}] \geq \rho, \qquad \forall n \geq 1.$$

Now, let us consider the sequence of random variables in $L^0_+(\Omega, \mathcal{F}_T, P)$

$$H^n = x_0 + \sum_{k=1}^n \hat{X}^k \mathbf{1}_{B^k}, \quad \text{where} \quad x_0 = \inf\{x > 0 : U(x) \geq 0\}.$$

Since $H^n \geq x_0$, we have $U(H^n) = U^+(H^n) \geq 0$.

For all $Q \in \mathcal{M}_e$, we have

$$E^Q[H^n] \leq x_0 + \sum_{k=1}^n E^Q[\hat{X}^k] \leq x_0 + nx,$$

since $\hat{X}^k \in \mathcal{C}(x)$. The characterisation on the Theorem (2.5.1) implies that $H^n \in \mathcal{C}(x_0 + nx)$. On the other hand,

$$\begin{aligned} E[U(H^n)] &= E[U^+(x_0 + \sum_{k=1}^n \hat{X}^k \mathbf{1}_{B^k})] \\ &\geq E[U^+(\sum_{k=1}^n \hat{X}^k \mathbf{1}_{B^k})] = \sum_{k=1}^n E[U^+(\hat{X}^k)\mathbf{1}_{B^k}] \geq \rho n, \end{aligned}$$

which hold true since U, U^+ are increasing. We thus deduce that

$$\begin{aligned} \limsup_{x\longrightarrow\infty} \frac{v(x)}{x} &= \limsup_{n\longrightarrow\infty} \frac{v(x_0 + nx)}{x_0 + nx} \\ &\geq \limsup_{n\longrightarrow\infty} \frac{E[U(H^n)]}{x_0 + nx} \\ &\geq \limsup_{n\longrightarrow\infty} \frac{\rho n}{x_0 + nx} = \frac{\rho}{x} > 0, \end{aligned}$$

which in fact contradicts (3.10). Hence, (3.22) holds true and $\hat{X}_T^x$ is solution to $v(x)$. From the strict concavity of U on $(0,\infty)$, if we consider $\hat{X}'^x_T$ and $\hat{X}''^x_T$ two different solutions of $v(x)$, we have for all $p \in [0,1]$, $p\hat{X}'^x_T + (1-p)\hat{X}''^x_T$ is also a solution of $v(x)$ and

$$U(p\hat{X}'^x_T + (1-p)\hat{X}''^x_T) > pU(\hat{X}'^x_T) + (1-p)U(\hat{X}''^x_T), \qquad \forall p \in [0,1].$$

By taking the expectation on both sides we have

$$E[U(p\hat{X}'^x_T + (1-p)\hat{X}''^x_T)] > pE[U(\hat{X}'^x_T)] + (1-p)E[U(\hat{X}''^x_T)], \qquad \forall p \in [0,1],$$

which implies

$$v(x) > pv(x) + (1-p)v(x) = v(x).$$

Thus, we get the contradiction. Therefore, the solution $\hat{X}_T^x$ to $v(x)$ is unique. ■

Problem (3.7) is a convex optimization problem under (infinite) linear constraints characterized by (2.38). A careful reader will notice that thanks to Theorem (3.2.3) we obtained existence and uniqueness of the solution to the value function v, however so far we still have not used any notion of duality and convexity techniques which is the main aim of this thesis. Because the solution of the problem (3.7) is not easy to find in the primal form, we will look at it in its dual form.

3.3 Resolution via the Dual Formulation

The static optimisation problem $v(x) = \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)]$ is stated as a concave maximisation problem on an infinite dimension in $L^0_+(\Omega, \mathcal{F}_T, P)$ subject to an infinity of linear constraints given by the dual characterisation (2.38) of $\mathcal{C}(x)$:

$$\begin{cases} v(x) = \sup\limits_{X_T \in L^0_+(\Omega,\mathcal{F}_T,P)} E[U(X_T)], \\ \text{subject to } \; E[Z_T X_T] \le x, \qquad \forall Z_T \in \mathcal{M}_e(S)\,. \end{cases} \tag{3.23}$$

Here and subsequently, a probability measure Q which is absolutely continuous with respect to P, i.e. $Q \ll P$, will be identified with its Radon-Nikodým density $Z_T = dQ/dP$ and to shorten notation, we write M_e instead of $M_e(S)$. The following questions arise in our mind:

- How may we calculate the value function $v(x)$?
- How may we calculate the optimal solution $\hat{X}_T \in \mathcal{C}(x)$ in (3.7), provided this solution exists by the theorem 3.2.3?

A well-known tool (see e.g. [Roc70]) to answer these questions is the passage to the conjugate function

$$\widetilde{U}(y) = \sup_{x>0}[U(x) - xy], \qquad y > 0. \tag{3.24}$$

The function $\widetilde{U}(y)$ is the Legendre-transform of the function $-U(-x)$ and we define the domain $\mathrm{dom}(\tilde{U}) = \{y > 0 : \widetilde{U}(x) < \infty\}$. It is well known (see e.g. [Roc70, ET76]), that if $U(x)$ satisfies the usual Inada conditions stated in (3.28) below, then $\widetilde{U}(y)$ is continuously differentiable, decreasing and strictly convex and satisfies the following conditions: $\widetilde{U}^{'}(0) = -\infty$, $\widetilde{U}^{'}(\infty) = 0$ and $\widetilde{U}(0) = U(\infty)$, $\widetilde{U}(\infty) = U(0)$. Let us denote by $I(\cdot)$ the (continuous, strictly decreasing) inverse of the marginal utility function U' on $(0, \infty)$; this function maps $(0, \infty)$ onto itself, and satisfies $I(0+) = \infty$, $I(\infty) = 0$. We have

$$I(y) := (U')^{-1}(y) = -\widetilde{U}'(y)\,, \quad 0 < y < \infty\,.$$

Furthermore, the supremum in (3.24) is attainable at $x = I(y)$ which means

$$\widetilde{U}(y) = U(I(y)) - yI(y)\,, \quad y > 0 \tag{3.25}$$

and the following bidual relation holds true:

$$U(x) = \inf_{y>0}[\widetilde{U}(y) + xy], \quad x > 0. \tag{3.26}$$

The infimum in 3.26 is attained at $y = U^{'}(x)$, i.e.

$$U(x) = \widetilde{U}(U'(x)) + xU'(x), \quad x > 0. \tag{3.27}$$

3.3.1 The Inada Condition for Utility

The Inada Condition is usually assumed for the production function in the economic growth model in order to make good economic sense. In words, we assume that the marginal utility U' of the first unit of consumption is arbitrarily large, and that the marginal utility of consumption goes to zero as per-capita consumption gets arbitrary large. It is stated as follow

Assumption 3.3.1 *[Inada Condition for Utility]*

$$\lim_{x\downarrow 0} U'(x) = +\infty, \qquad \lim_{x\uparrow +\infty} U'(x) = 0. \tag{3.28}$$

This condition prevents the optimal solution of zero consumption (or holding) in maximisation problems, which is useful to ensure a non-trivial positive solution. However, the Inada condition is not satisfied with a very common utility function, which is the exponential utility, $U(x) = 1 - e^{-x}$. This utility satisfies another condition instead of the Inada condition which is given by

$$\lim_{x\downarrow -\infty} U'(x) = +\infty, \qquad \lim_{x\uparrow +\infty} U'(x) = 0.$$

This allows us to have the optimal solution in the non-positive consumption (or holding), so this condition is weaker than the Inada Condition.

Example 3.3.2 *Typical examples of utility functions satisfying the Inada Condition (3.3.1), and their conjugate functions are*

$$\begin{array}{rcl} U(x) & = & \ln(x), \qquad \widetilde{U}(y) = -\ln(y) - 1 \\ U(x) & = & \dfrac{x^p}{p}, \quad 0 < p < 1\,, \qquad \widetilde{U}(y) = \dfrac{y^{-q}}{q}, \;\; q = \dfrac{p}{1-p} \\ U(x) & = & -\dfrac{e^{-ax}}{a}, \quad a > 0\,, \qquad \widetilde{U}(y) = \dfrac{y}{a}(\ln(y) - 1). \end{array}$$

3.3.2 Saddlepoint Problem

Using the definition of U in (3.26) we have the following inequality:

$$U(x) \leq \widetilde{U}(y) + xy\,, \quad \forall\, x > 0\,, y > 0\,. \tag{3.29}$$

Therefore, for all $x > 0$, $y > 0$, $X_T \in \mathcal{C}(x)$, $Z_T \in \mathcal{M}_e$,

$$U(X_T) \;\; \leq \;\; \widetilde{U}(yZ_T) + yZ_T X_T\,. \tag{3.30}$$

By taking the expectation in both sides of (3.30) and thanks to the dual characterisation (2.38), the following inequalities hold true:

$$\begin{array}{rcl} E[U(X_T)] & \leq & E[\widetilde{U}(yZ_T)] + E[yZ_T X_T] \\ & \leq & E[\widetilde{U}(yZ_T)] + yx. \end{array} \tag{3.31}$$

Inequality (3.31) is the point of departure of the dual resolution approach. Then the dual problem to $v(x)$ in (3.7) is introduced as follows:

$$\tilde{v}(y) = \inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)], \qquad y > 0. \tag{3.32}$$

The inequality (3.31) implies that for all $x > 0$

$$\begin{aligned} v(x) &= \sup_{X_T \in \mathcal{C}(x)} E[U(X_T)] \\ &\leq \inf_{y>0}\{\tilde{v}(y) + xy\} = \inf_{Z_T \in \mathcal{M}_e, y>0}\{E[\widetilde{U}(yZ_T)] + xy\}. \end{aligned} \tag{3.33}$$

The main dual resolution method to the primal problem $v(x)$ consists in the following procedure: prove the existence of a $(\hat{y}, \hat{Z}_T)$ (depending on x) solution to the dual problem in the right-hand side of (3.33). This is equivalent to showing the existence of $\hat{y} > 0$ attaining the minimum of $\tilde{v}(y) + xy$, and to obtain the existence of a $\hat{Z}_T$ solution to the dual problem $\tilde{v}(\hat{y})$. We then set

$$\hat{X}_T^x = I(\hat{y}\hat{Z}_T) \quad i.e. \quad U'(\hat{X}_T^x) = \hat{y}\hat{Z}_T\,.$$

From the first-order optimality conditions on $\hat{y}$ and $\hat{Z}_T$, we shall see that this implies

$$\hat{X}_T^x \in \mathcal{C}(x) \quad \text{and} \quad E[\hat{X}_T^x \hat{Z}_T] = x\,.$$

From (3.27), we then obtain

$$E[U(\hat{X}_T^x)] = E[\widetilde{U}(\hat{y}\hat{Z}_T)] + x\hat{y}\,.$$

which proves, recalling (3.33), that

$$E[U(\hat{X}_T^x)] \leq v(x) \leq E[\widetilde{U}(\hat{y}\hat{Z}_T)] + x\hat{y} = E[U(\hat{X}_T^x)]\,,$$

i.e. $\hat{X}_T^x$ is solution to $v(x)$. Furthermore, the conjugate duality relations on the primal and dual value functions hold true:

$$v(x) = \inf_{y>0}\{\tilde{v}(y) + xy\} = \tilde{v}(\hat{y}) + x\hat{y}\,.$$

Before we discuss in the next Subsection 3.3.3 the technical difficulties arising in this dual approach, we can already at this stage provide some sufficient conditions ensuring that assumption (3.10) holds, thus the existence (and uniqueness) of a solution to the primal problem $v(x)$ is secured.

Remark 3.3.3 *Let us assume that*

$$\tilde{v}(y) < \infty\,, \quad \forall\, y > 0. \tag{3.34}$$

From inequality (3.33), one easily show that condition (3.10) is satisfied. The condition (3.34) is clearly verified once we have

$$\forall\, y > 0\,, \quad \exists\, Z_T \in \mathcal{M}_e \ \textit{ such that } \ E[\widetilde{U}(yZ_T)] < \infty\,. \tag{3.35}$$

A condition bearing directly on U and ensuring (3.35) is: there exist $p \in (0,1)$, positive constants k_1, k_2 and $Z_T \in \mathcal{M}_e$ with the following properties:

$$U^+(x) \leq k_1 x^p + k_2\,, \quad \forall\, x > 0\,, \tag{3.36}$$

$$E[Z_T^{-q}] < \infty\,, \quad \textit{where} \quad q = \frac{p}{1-p} > 0\,. \tag{3.37}$$

Indeed, in this case, we have

$$\widetilde{U}(y) \leq \sup_{x>0}[k_1 x^p - xy] + k_2 = (k_1 p)^{\frac{1}{1-p}} \frac{y^{-q}}{q} + k_2\,, \quad \forall\, y > 0\,,$$

and (3.35) is obviously satisfied. As indicated later on, a weaker condition (in fact a minimal condition) on U, namely the Reasonable Asymptotic Elasticity Assumption 3.3.7, will ensure condition (3.34).

3.3.3 Dual Space Variables

Let us now study in more details the dual approach sketched out in a rather formal way in Subsection 3.3.2. The existence of a solution to the dual problem $\tilde{v}(y)$, $y > 0$ is the matter that we need to deal with carefully. The set $\mathcal{M}_e$ on which the optimisation $\tilde{v}(y) = \inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)]$, $y > 0$ is attained is naturally included in $L^1(\Omega, \mathcal{F}_T, P)$, however in the infinite dimensional space L^1 there is no compactness result. The Komlos Theorem A.3.1 asserts that from any bounded sequence $(Z^n)_n$ in L^1, there are a random variable $\hat{Z} \in L^1$ and a subsequence $(Z^{n_k})_{k\geq 1}$ Cesaro-convergent[†] almost surely to $\hat{Z}$. However, this convergence does not hold true in general in L^1. Therefore, we do not have in general that $E[\frac{1}{k}\sum_{j=1}^{k} Z^{n_j}] \longrightarrow E[\hat{Z}]$ when we let k to infinity. In our problem (3.32), the maximising

[†] Recall that a sequence (a_m) is called Cesaro-convergent if the sequence of arithmetic means $\bar{a}_m = m^{-1}\sum_{k\leq m} a_k$ converges.

sequence of probability measures $(Z^n)_{n\geq 1}$ in $\mathcal{M}_e$ (which satisfies $E[Z^n]=1$) for $\tilde{v}(y)$ does not necessarily converge to a probability measure $\hat{Z}_T$: in general, we have $E[\hat{Z}_T] < 1$. In fact, the space $L^0_+(\Omega, \mathcal{F}_T, P)$ in which the primal variables $X_T \in \mathcal{C}(x)$ vary, is not in appropriate duality with $L^1(\Omega, \mathcal{F}_T, P)$. Thus, it is more natural to let the dual variables also vary in $L^0_+(\Omega, \mathcal{F}_T, P)$. We then "enlarge" the set $\mathcal{M}_e$ in $L^0_+(\Omega, \mathcal{F}_T, P)$ as follows. We define $\mathcal{D}$ as the convex, solid and closed envelope of $\mathcal{M}_e$ in $L^0_+(\Omega, \mathcal{F}_T, P)$ i.e. the smallest convex, solid and closed subset in $L^0_+(\Omega, \mathcal{F}_T, P)$ containing $\mathcal{M}_e$. It is easy to show (using, e.g. Theorem A.3.2), that $\mathcal{D}$ equals

$$\mathcal{D} = \{Y_T \in L^0_+(\Omega, \mathcal{F}_T, P) \,:\, \exists (Z^n)_{n\geq 1} \in \mathcal{M}_e, \quad Y_T \leq \lim_{n\to\infty} Z^n\}, \tag{3.38}$$

where the limit $\lim_{n\to\infty} Z^n$ is understood in the sense of almost surely convergence.

From the dual space characterization (2.38) and Fatou Lemma, we conclude that the set $\mathcal{C}(x)$ is also written in duality relation with $\mathcal{D}$: for $X_T \in L^0_+(\Omega, \mathcal{F}_T, P)$,

$$X_T \in \mathcal{C}(x) \Longleftrightarrow E[Y_T X_T] \leq x\,, \qquad \forall Y_T \in \mathcal{D}. \tag{3.39}$$

Let us summarize the above properties of $\mathcal{D}$ in the following proposition:

Proposition 3.3.4 *The set $\mathcal{D}$ defined in (3.38) is a subset of $L^0_+(\Omega, \mathcal{F}_T, P)$ which is convex, solid and closed in the topology of convergence in measure. Moreover, for any $x > 0$, we have the following polarity relation*

$$X_T \in \mathcal{C}(x) \Longleftrightarrow E[Y_T X_T] \leq x\,, \qquad \forall\, Y_T \in \mathcal{D}\,, \tag{3.40}$$

and

$$Y_T \in \mathcal{D} \Longleftrightarrow E[Y_T X_T] \leq x\,, \qquad \forall\, X_T \in \mathcal{C}(x)\,. \tag{3.41}$$

We then consider the dual problem

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E[\widetilde{U}(yY_T)], \quad y > 0 \qquad \textsf{(Dual problem)}. \tag{3.42}$$

We shall see below that this definition is consistent with the one in (3.32), i.e. for $y > 0$, we have

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E[\widetilde{U}(yY_T)] = \inf_{Y_T \in \mathcal{M}_e} E[\widetilde{U}(yY_T)]. \tag{3.43}$$

In other words, the infimum in $\tilde{v}(y)$ coincides when it is taken over $\mathcal{M}_e$ or $\mathcal{D}$.

The last ingredient needed before we state and proof Kramkov-Schachermayer Theorem, is to introduce and illustrate the asymptotic elasticity condition.

3.3.4 The Asymptotic Elasticity Condition

Definition 3.3.5 *For a utility function $U : (0,\infty) \to \mathbb{R}$, the "asymptotic elasticity"(at ∞) is defined as*

$$AE(U) := \limsup_{x\to\infty} \frac{xU'(x)}{U(x)}. \tag{3.44}$$

The economic interpretation of the asymptotic elasticity $AE(U)$ is the limit of the ratio between the marginal utility $U'(x)$ and the average utility $\frac{U(x)}{x}$, as $x \to \infty$.

We recall the following lemma from [KS98], from which it may be concluded that, for any concave function U with the property that the right hand side is well defined, we always have that $AE(U) \leq 1$.

Lemma 3.3.6 *For a strictly concave, increasing, real-valued function U the definition of the asymptotic elasticity $AE(U)$ makes sense and, depending on $U(\infty) = \lim_{x\to\infty} U(x)$, takes its values in the following sets:*

1. *For $U(\infty) = \infty$ we have $AE(U) \in [0,1]$,*
2. *For $0 < U(\infty) < \infty$ we have $AE(U) = 0$,*
3. *For $-\infty < U(\infty) \leq 0$ we have $AE(U) \in [-\infty, 0]$.*

Proof. [KS98], Lemma 6.1, page 943 for the proof. ∎

Classical examples (and counter-examples) of such utility functions are:

- $U(x) = \ln x$, for which $AE(U) = 0$
- $U(x) = \frac{x^p}{p}$, $0 < p < 1$, for which $AE(U) = p$
- $U(x) = \frac{x}{\ln x}$, for x large enough, for which $AE(U) = 1$

Likewise as in Definition 3.3.5 we define the asymptotic elasticity $AE(v)$ of the value function v by

$$AE(v) := \limsup_{x\to\infty} \frac{xv'(x)}{v(x)}. \tag{3.45}$$

Economic intuition suggests that the marginal utility $U'(x)$ should be substantially smaller than the average utility $\frac{U(x)}{x}$, as $x \to \infty$. This leads us to the following reasonable asymptotic elasticity Assumption:

Assumption 3.3.7 *The utility function $U : (0,\infty) \to \mathbb{R}$ has asymptotic elasticity at infinity strictly less than 1, i.e.*

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1. \tag{3.46}$$

We shall see that Assumption 3.3.7 is the crucial condition for existence of the optimal solution to the maximization problem (3.7).

We state below the key result in optimisation via the convex duality approach.

3.4 The Kramkov-Schachermayer Theorem

The following theorem due to Kramkov and Schachermayer states that under the (3.1.1), (3.1.2), (3.3.1) and (3.3.7)) assumptions on the utility function U and the value function v, the duality theory "works" well in this context. In fact, Assumption 3.3.7 of reasonable asymptotic elasticity $AE(U) < 1$ is minimal (i.e. necessary and sufficient) and cannot be weakened in the sense that one can find counter-examples of continuous price processes S for which the value function $\tilde{v}(y)$ is not finite for all y and there does not exist a solution to the primal problem $v(x)$, whenever $AE(U) = 1$ (we refer the interested reader to the seminal paper by Kramkov and Schachermayer [KS99] for more details).

Theorem 3.4.1 (Kramkov-Schachermayer) *Let U be a utility function satisfying assumptions (3.1.1), (3.1.2), (3.3.1) and (3.3.7). Then, the following assertions hold:*

(1) *The function v is finite, continuously differentiable, strictly concave on $(0,\infty)$, and there exists a unique solution $\hat{X}_T^x \in \mathcal{C}(x)$ to $v(x)$ for all $x > 0$. Moreover, the function v' is strictly decreasing and satisfies*

$$v'(\infty) = \lim_{x\to\infty} v'(x) = 0\,, \quad \textit{and} \quad v'(0) = \lim_{x\to 0} v'(x) = \infty\,.$$

(2) *The function $\tilde{v}$ is finite, continuously differentiable, strictly convex on $(0,\infty)$, and there exists a unique solution $\hat{Y}_T^y \in \mathcal{D}$ to $\tilde{v}(y)$ for all $y > 0$. Moreover, the function $-\tilde{v}'$ is strictly decreasing and satisfies*

$$\tilde{v}'(\infty) := \lim_{y\longrightarrow\infty} \tilde{v}'(y) = 0 \quad \textit{and} \quad -\tilde{v}'(0) = \lim_{y\to 0} -\tilde{v}'(y) = \infty\,.$$

(3) (i) *For all $x > 0$, we have*

$$\hat{X}_T^x = I(\hat{y}\hat{Y}_T), \;\; i.e. \;\; U^{'}(\hat{X}_T^x) = \hat{y}\hat{Y}_T, \tag{3.47}$$

where $\hat{Y}_T \in \mathcal{D}$ is the solution of $\tilde{v}(\hat{y})$ with $\hat{y} = v'(x)$ the solution to $argmin_{y>0}[\tilde{v}(y) + xy]$ and satisfying

$$E[\hat{Y}_T \hat{X}_T^x] = x \,, \quad i.e. \quad v'(x) = E[\frac{\hat{X}_T^x U^{'}(\hat{X}_T^x)}{x}] \,. \tag{3.48}$$

(ii) *We have the conjugate duality relations*

$$\begin{cases} v(x) = \min_{y>0}\{\tilde{v}(y) + xy\}, & \forall \, x > 0 \\ \\ \tilde{v}(y) = \max_{x>0}\{v(x) - xy\}, & \forall \, y > 0. \end{cases} \tag{3.49}$$

(4) *Furthermore, if there exists $y > 0$ such that $\inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)] < \infty$, then*

$$\tilde{v}(y) = \inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)] = \inf_{Y_T \in \mathcal{D}} E[\widetilde{U}(yY_T)] \tag{3.50}$$

Remark 3.4.2 *Denote by $\hat{X}_t^x = x + \int_0^t \hat{\alpha}_u dS_u$, $0 \leq t \leq T$, the optimal wealth process associated to the primal problem $v(x)$. (there is no ambiguity of notation at the expiration time T since the optima portfolio $\hat{X}_T^x$ as shown in equation (3.47) is the one that perfectly replicates the inverse of marginal utility, evaluated (up to a multiplicative constant) at the optimal solution $\hat{Y}_T$ in the dual problem, i.e. $x + \int_0^T \hat{\alpha}_u dS_u = \hat{X}_T^x = I(\hat{y}\hat{Y}_T)$ is indeed the solution in $\mathcal{C}(x)$ to $v(x)$). The process $\hat{X}^x$ is equal (up to a càd-làg modification) to*

$$\hat{X}_t^x = \operatorname*{ess\,sup}_{Q \in \mathcal{M}_e} E^Q[I(\hat{y}\hat{Y}_T)|\mathcal{F}_t], \quad 0 \leq t \leq T \,.$$

and the optimal control $\hat{\alpha}$ is determined from the optional decomposition of $\hat{X}^x$. In the case where the dual problem $\tilde{v}(\hat{y})$ admits a solution $\hat{Z}_T$ in $\mathcal{M}_e$ with corresponding probability measure $\hat{Q}$, then the process $\hat{X}^x$ is a nonnegative local martingale under $\hat{Q}$, hence a $\hat{Q}$-supermartingale such that $E^{\hat{Q}}[\hat{X}_T^x] = x = E^{\hat{Q}}[\hat{X}_0^x]$ by (3.48). It follows that $\hat{X}^x$ is a $\hat{Q}$-martingale (this fact is also a direct application of Proposition 2.4.6), which is consequently written as

$$\hat{X}_t^x = E^{\hat{Q}}[I(\hat{y}\hat{Z}_T)|\mathcal{F}_t], \quad 0 \leq t \leq T \,.$$

Remark 3.4.3 *The assertion* (4) *still holds if we drop the condition that there exists* $y > 0$ *such that* $\inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)] < \infty$ *(see Proposition 3.2 in [KS99]).*

The remaining of this Chapter provides a detailed proof of Kramkov-Schachermayer Theorem 3.4.1. We divide the long proof in several lemmas and propositions of independent interests, where the required assumptions are clearly indicate for each step of the proof.

Lemma 3.4.4 *Let* U *be a utility function satisfying assumptions (3.1.1) and (3.3.1). Then, for all* $y > 0$*, the family* $\{\widetilde{U}^-(yY_T), Y_T \in \mathcal{D}\}$ *is uniformly integrable.*

Proof. Since $\widetilde{U}$ is strictly decreasing from $(0, \infty)$ into $(\widetilde{U}(0), \widetilde{U}(\infty))$. If $\tilde{U}(\infty) < -\infty$, the continuous function $\widetilde{U}$ is bounded and we have

$$\tilde{U}^- \leq |\widetilde{U}| \leq |\widetilde{U}(0)| \vee |\widetilde{U}(\infty)|.$$

Therefore the family of random variables $\{\tilde{U}^-(yY_T), Y_T \in \mathcal{D}\}$ is uniformly integrable.

Let us consider the case when $\widetilde{U}(\infty) = -\infty$ and let ϕ be the inverse function of $-\widetilde{U}$. It is clear that ϕ is strictly increasing function from $(-\tilde{U}(0), \infty)$ into $(0, \infty)$. Recall that $\widetilde{U}(0) = U(\infty) > 0$, and so ϕ is well-defined on $[0, \infty)$. Let us denote $X^- = \max(-X, 0)$. The following inequality holds true

$$\begin{aligned}
E[\phi(X^-)] &= E[\phi(X^-)\mathbf{1}_{(X \geq 0)}] + E[\phi(X^-)\mathbf{1}_{(X<0)}] \\
&= E[\phi(0)\mathbf{1}_{(X \geq 0)}] + E[\phi(-X)\mathbf{1}_{(X<0)}] \\
&= E[\phi(-X)\mathbf{1}_{(X<0)}] + \phi(0)E[\mathbf{1}_{(X \geq 0)}] \\
&\leq E[\phi(-X)] + \phi(0)P(X \geq 0) \\
&\leq E[\phi(-X)] + \phi(0), \qquad (3.51)
\end{aligned}$$

since ϕ is positive.

Now, using (3.51) and $\phi = (-\tilde{U})^{-1}$, we have for all $y > 0$,

$$\begin{aligned}
E[\phi(\widetilde{U}^-(yY_T))] &\leq E[\phi(-\widetilde{U}(yY_T))] + \phi(0) = E[yY_T] + \phi(0) \\
&\leq y + \phi(0),
\end{aligned}$$

since $X_T = 1 \in \mathcal{C}(1)$. In the other hand, by changing the variable $y = \phi(x)$ and using the l'Hôpital rule, we have from the Inada condition (3.28) that

$$\lim_{x \to \infty} \frac{\phi(x)}{x} = \lim_{y \to \infty} \frac{y}{-\widetilde{U}(y)} = \lim_{x \to \infty} \frac{1}{I(y)} = \infty.$$

Therefore, by the Theorem of la Vallé-Poussin we conclude that the family $\{\widetilde{U}^-(yY_T), Y_T \in \mathcal{D}\}$ is uniformly integrable. ∎

Lemma 3.4.5 *The following equality holds*

$$\sup_{x>0} \sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} = \sup_{x>0} \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\}, \tag{3.52}$$

Proof. Let us fix $x > 0$, (3.39) implies that

$$X_T \in \mathcal{C}(x) \quad \Leftrightarrow \quad E[X_TY_T] \leq x, \quad ,\forall Y_T \in \mathcal{D} \tag{3.53}$$

Therefore, we have for $y > 0$

$$-xy \leq \inf_{X_T \in \mathcal{C}(x)} \{-E[yX_TY_T]\}, \quad \forall Y_T \in \mathcal{D} \tag{3.54}$$

i.e.

$$-xy \leq -\sup_{X_T \in \mathcal{C}(x)} \{E[yX_TY_T]\}, \quad \forall Y_T \in \mathcal{D} \tag{3.55}$$

We want to show that for any $y > 0$

$$\sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} = \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\}, .$$

We start first by proving that

$$\sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} \leq \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\}, .$$

Indeed, we have

$$\begin{aligned} \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\} &= \sup_{X_T \in \mathcal{C}(x)} \{E[U(X_T)] + \inf_{Y_T \in \mathcal{D}} E[-yX_TY_T]\} \\ &= \sup_{X_T \in \mathcal{C}(x)} \{E[U(X_T)] - \sup_{Y_T \in \mathcal{D}} E[yX_TY_T]\} \\ &\geq \sup_{X_T \in \mathcal{C}(x)} \{E[U(X_T) - xy]\} \end{aligned}$$

which is true thanks to (3.55) and the fact that $\inf(-f) = -\sup(f)$. Thus, we have prove that, for all $x > 0$

$$\sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} \leq \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\},$$

which implies

$$\sup_{x>0} \sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} \leq \sup_{x>0} \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\}, \tag{3.56}$$

Now we want to prove the converse inequality, i.e.

$$\sup_{X_T \in \mathcal{C}(x)} E\{U(X_T) - xy\} \geq \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\}.$$

For a given $X_T \in \mathcal{C}(x)$, (3.53) implies

$$\sup_{Y_T \in \mathcal{D}} E\{yX_TY_T\} \quad \leq \quad xy, \quad \forall y > 0$$

which implies that, there exist z such that $0 < z \leq x$ and for $y > 0$, we have

$$\sup_{Y_T \in \mathcal{D}} E\{yX_TY_T\} = yz \leq xy,$$

therefore

$$\sup_{Y_T \in \mathcal{D}} E\{X_TY_T\} \leq z$$

i.e. $X_T \in \mathcal{C}(z)$, and thus,

$$\begin{aligned}
\inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\} &= E\{U(X_T)\} - \sup_{Y_T \in \mathcal{D}} E[yX_TY_T] \\
&= E\{U(X_T)\} - yz \\
&\leq \sup_{\widetilde{X}_T \in \mathcal{C}(z)} E\{U(\widetilde{X}_T) - yz\} \\
&\leq \sup_{0<z\leq x} [\sup_{\widetilde{X}_T \in \mathcal{C}(z)} E\{U(\widetilde{X}_T) - yz\}].
\end{aligned} \tag{3.57}$$

In other words, for an arbitrary $X_T \in \mathcal{C}(x)$, we have

$$\inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\} \leq \sup_{0<z\leq x} [\sup_{\tilde{X}_T \in \mathcal{C}(z)} E\{U(\tilde{X}_T) - yz\}].$$

where the R.H.S is independent of X_T, therefore we have the following

$$\sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} E\{U(X_T) - yX_TY_T\} \leq \sup_{0<z\leq x} \sup_{\tilde{X}_T \in \mathcal{C}(z)} E\{U(\widetilde{X}_T) - yz\}, \tag{3.58}$$

which implies

$$\begin{aligned}
\sup_{x>0}\sup_{X_T\in\mathcal{C}(x)}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\} &\leq \sup_{x>0}\sup_{0<z\leq x}[\sup_{\tilde{X}_T\in\mathcal{C}(z)} E\{U(\tilde{X}_T)-yz\}] \\
&= \sup_{z>0}\sup_{\widetilde{X}_T\in\mathcal{C}(z)} E\{U(\widetilde{X}_T)-yz\}] \\
&= \sup_{x>0}\sup_{X_T\in\mathcal{C}(x)} E\{U(X_T)-yx\},
\end{aligned}$$

The above nequality (3.59) shows that

$$\sup_{x>0}\sup_{X_T\in\mathcal{C}(x)} E\{U(X_T)-xy\} \geq \sup_{x>0}\sup_{X_T\in\mathcal{C}(x)}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\}. \tag{3.59}$$

Therefore, (3.56) and (3.59) prove that equality in (3.52) holds true.

■

Lemma 3.4.6 *The following equality holds*

$$\sup_{X_T\in L^0_+}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\} = \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\}. \tag{3.60}$$

Proof. We have

$$\sup_{X_T\in L^0_+}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\} \leq \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\}, \tag{3.61}$$

which is true, since for any $x>0$, we have $X\equiv x\in\mathcal{C}(x)$ and therefore $\cup_{n=1}^{\infty}\mathcal{B}_n = L^0_+ \subseteq \cup_{x>0}\mathcal{C}(x)$.

Conversely, let us fix $\varepsilon>0$, by the definition of the supremun, we have, there exist $X_\varepsilon\in\cup_{x>0}\mathcal{C}(x)$ such that

$$\inf_{Y_T\in\mathcal{D}} E\{U(X_\varepsilon)-yX_\varepsilon Y_T\} \geq \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)}\inf_{Y_T\in\mathcal{D}} E\{U(X_T)-yX_TY_T\}-\varepsilon. \tag{3.62}$$

Let us consider the sequence $X_k = X_\varepsilon\wedge k = \min(k, X_\varepsilon)\leq k$. We have $X_k\in\mathcal{B}_k\subseteq L^0_+$ is an increasingly sequence in $\mathcal{B}_k$ and $X_k\uparrow X_\varepsilon$ a.s., we have

$$\begin{aligned}
\inf_{Y_T\in\mathcal{D}} E\{U(X_k)-yX_kY_T\} &= E\{U(X_k)\}-\sup_{Y_T\in\mathcal{D}} E[yX_kY_T] \\
&\geq E\{U(X_k)\}-\sup_{Y_T\in\mathcal{D}} E[yX_\varepsilon Y_T],
\end{aligned} \tag{3.63}$$

since $X_k\leq X_\varepsilon$ by construction.

$$\inf_{Y_T\in\mathcal{D}} E\{U(X_k) - yX_kY_T\} \geq E\{U(X_k)\} - \sup_{Y_T\in\mathcal{D}} E[yX_\varepsilon Y_T], \tag{3.64}$$

which implies

$$\begin{aligned} \liminf_k \inf_{Y_T\in\mathcal{D}} E\{U(X_k) - yX_kY_T\} &\geq \liminf_k E\{U(X_k)\} - \sup_{Y_T\in\mathcal{D}} E[yX_\varepsilon Y_T] \\ &= E[U(X_\varepsilon)] - \sup_{Y_T\in\mathcal{D}} E[yX_\varepsilon Y_T] \\ &= \inf_{Y_T\in\mathcal{D}} E[U(X_\varepsilon) - yX_\varepsilon Y_T] \\ &\geq \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] - \varepsilon, \end{aligned} \tag{3.65}$$

since, by the convergence monotone theorem $\liminf_k E\{U(X_k)\} = E[U(X_\varepsilon)]$ and the definition of X_ε. Inequality (3.65) implies there exist an integer k such that $X_k \in L^0_+$ and

$$\inf_{Y_T\in\mathcal{D}} E[U(X_k) - yX_kY_T] \geq \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] - \varepsilon. \tag{3.66}$$

Therefore,

$$\sup_{X_T\in L^0_+} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] \geq \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \inf_{Y_T\in\mathcal{D}} E[U(X_T) - yX_TY_T] - \varepsilon, \tag{3.67}$$

which hold true by definition of limit inf. Thus, by taking $\varepsilon \longrightarrow 0$, we obtain the result.

■

Example 3.4.7 *A simple illustrative example will show the relevance of the above lemma. In an arbitrage-free and complete financial market, a call option is priced under the unique equivalent martingale measure*

$$x = E^{\mathbb{Q}}[(S_T - K)^+] < \infty\,.$$

Starting with the initial endowment x computed by the equation above, we can find a non-negative self-financing portfolio that exactly replicates the call option payoff at expiration time T,

$$X_T = x + \int_0^T \xi_t dS_t = (S_T - K)^+\,.$$

Hence by definition

$$X_T \in \mathcal{C}(x)\,.$$

However, because S_T could be unbounded, we may have

$$X_T \notin \mathcal{B}_n\,, \quad \text{for any} \quad n \geq 1\,.$$

Proposition 3.4.8 *(Conjugate duality relations)*

Let U be a utility function satisfying (3.1.1), (3.1.2) and (3.3.1). Then, the value functions v and $\tilde{v}$ are linked by the duality relations:

$$v(x) = \inf_{y>0}\{\tilde{v}(y) + xy\}, \qquad x > 0 \tag{3.68}$$

$$\tilde{v}(y) = \sup_{x>0}\{v(x) - xy\}, \qquad y > 0. \tag{3.69}$$

Proof. Using 3.33 we have for all $x > 0$

$$v(x) \leq \inf_{y>0}\{\tilde{v}(y) + xy\}, \qquad \forall y > 0, \tag{3.70}$$

which implies that for all $x > 0$

$$v(x) - xy \leq \tilde{v}(y), \qquad \forall y > 0,$$

which implies that

$$\sup_{x>0}\{v(x) - xy\} \leq \tilde{v}(y), \qquad \forall y > 0. \tag{3.71}$$

Fix some $y > 0$. To show the second equation in the system of 3.68, we assume without loss in generality that $\sup_{x>0}\{v(x) - xy\} < \infty$. For all $n > 0$ we consider the set

$$\mathcal{B}_n = \{X_T \in L^0_+(\Omega, \mathcal{F}_T, P) : X_T \leq n, a.s.\}.$$

$\mathcal{B}_n$ is compact in L^∞ for the weak topology $\sigma(L^\infty, L^1)$ and since $\mathcal{D}$ is a convex, closed subset of $L^1(\Omega, \mathcal{F}_T, P)$, we apply the Minimax Theorem (see B.1.3) and we have

$$\sup_{X_T\in\mathcal{B}_n} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} = \inf_{Y_T\in\mathcal{D}} \sup_{X_T\in\mathcal{B}_n} E\{U(X_T) - yX_TY_T\}, \tag{3.72}$$

for all $n > 0$ and $y > 0$.

Using the definition of $v(x)$, we have

$$\sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} E\{U(X_T) - xy\} = \sup_{x>0}\{v(x) - xy\} \tag{3.73}$$

From the duality relation 3.39 between $\mathcal{D}$ and $\mathcal{C}(x)$ we shall prove that

$$\lim_{n\longrightarrow\infty} \sup_{X_T\in\mathcal{B}_n} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} = \sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} E\{U(X_T) - xy\}. \tag{3.74}$$

We remark that $(\mathcal{B}_n)$ is increasing and $\cup_{n=1}^{\infty}\mathcal{B}_n = L_+^0(\Omega, \mathcal{F}_T, P)$ and also that for $0, x_1 < x_2$, we have $\mathcal{C}(x_1) \subset \mathcal{C}(x_2)$. We also remark that, for a given function ϕ, we have

$$\lim_{n\longrightarrow\infty} \sup_{X_T\in\mathcal{B}_n} \phi(X_T) = \sup_{X_T\in\cup_{n=1}^{\infty}\mathcal{B}_n} \phi(X_T) \tag{3.75}$$

$$\sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} \phi(X_T) = \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \phi(X_T). \tag{3.76}$$

Therefore, the right hand side of (3.74) becomes

$$\sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} E\{U(X_T) - xy\} = \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} E\{U(X_T) - xy\}, \tag{3.77}$$

By combining the result of Lemma 3.4.6,

$$\sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} = \sup_{X_T\in L_+^0} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\}.$$

and of Lemma 3.4.5

$$\sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} = \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} E\{U(X_T) - xy\},$$

We get the desired equality:

$$\sup_{X_T\in L_+^0} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} = \sup_{X_T\in\cup_{x>0}\mathcal{C}(x)} E\{U(X_T) - xy\}. \tag{3.78}$$

In other words, we have proved that

$$\begin{aligned}
\lim_{n\to\infty} \sup_{X_T\in\mathcal{B}_n} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} &= \sup_{X_T\in L_+^0} \inf_{Y_T\in\mathcal{D}} E\{U(X_T) - yX_TY_T\} \\
&= \sup_{x>0} \sup_{X_T\in\mathcal{C}(x)} E\{U(X_T) - xy\} \\
&= \sup_{x>0}\{v(x) - xy\}
\end{aligned}$$

Equivalently, thanks to the Minimax Theorem (see B.1.3)

$$\lim_{n\to\infty} \inf_{Y_T\in\mathcal{D}} \sup_{X_T\in\mathcal{B}_n} E\{U(X_T) - yX_TY_T\} = \sup_{x>0}\{v(x) - xy\} \tag{3.79}$$

Now, let us come back to our next step of the proof of (3.68) by defining

$$\tilde{U}_n(y) := \sup_{0<x\leq n} [U(x) - xy], \qquad y > 0,$$

we have

$$\inf_{Y_T \in \mathcal{D}} \sup_{X_T \in \mathcal{B}_n} E\{U(X_T) - yX_TY_T\} = \inf_{Y_T \in \mathcal{D}} E\{\tilde{U}_n(yY_T)\} =: \tilde{v}_n(y),$$

so that, by using (3.79) we obtain

$$\lim_{n\to\infty} \tilde{v}_n(y) = \sup_{x>0}[v(x) - xy] < \infty.$$

Thus, to obtain the second equation in (3.68), we must show that

$$\lim_{n\to\infty} \tilde{v}_n(y) = \tilde{v}(y). \tag{3.80}$$

We can see clearly that for a fixed y, the sequence $\tilde{v}_n(y)$ is increasing in n and

$$\tilde{v}_n(y) = \inf_{Y_T \in \mathcal{D}} E\{\tilde{U}_n(yY_T)\} \leq \inf_{Y_T \in \mathcal{D}} E\{\tilde{U}(yY_T)\} = \tilde{v}(y) \quad \forall n \geq 1.$$

Therefore, we have $\lim_{n\to\infty} \tilde{v}_n(y) \leq \tilde{v}(y)$. Let us prove the converse, for this let us consider $(Y_T^n)_{n\geq 1}$ a minimizing sequences in $\mathcal{D}$ of $\tilde{v}_n(y)$: we have

$$\lim_{n\to\infty} E\{\tilde{U}_n(yY_T^n)\} = \lim_{n\to\infty} \tilde{v}_n(y) < \infty.$$

In fact, by definition, we have

$$\tilde{v}_n(y) = \inf_{Y_T \in \mathcal{D}} E[\tilde{U}_n(yY_T)],$$

which implies there exists $Y_T^n \in \mathcal{D}$ such that

$$\tilde{v}_n(y) \leq E[\tilde{U}_n(yY_T^n)] \leq \tilde{v}_n(y) + \frac{1}{n},$$

which implies

$$\lim_{n\to\infty} \tilde{v}_n(y) \leq \lim_{n\to\infty} E[\tilde{U}_n(yY_T^n)] \leq \lim_{n\to\infty} \tilde{v}_n(y).$$

From the compactness theorem in $L^0_+(\Omega, \mathcal{F}_T, P)$, we can find a convex combination $\hat{Y}^n \in conv(Y^n, Y^{n+1}, \cdots)^{\ddagger}$, which is still lying in the convex set $\mathcal{D}$, and converges a.s. to a non-negative random variable Y_T. Since $\mathcal{D}$ is closed for the convergence in measure, we have $Y_T \in \mathcal{D}$.

$^{\ddagger}\hat{Y}^n \in conv(Y^n, Y^{n+1}, \cdots)$, means that for any $n \geq 1$, $\hat{Y}^n = \sum_{j=0}^{k(n)} a_j^n Y^{n+j}$, with positive constant such that $\sum_{j=0}^{k(n)} a_j^n = 1$.

We have $\hat{Y}^n \in conv(Y_T^n, Y_T^{n+1}, \cdots)$,$\hat{Y}^n \in \mathcal{D}$, and $\hat{Y}^n \to Y_T$, implies

$$\hat{Y}^n = \sum_{k\geq n} a_k Y_T^k, \qquad \sum_{k\geq n} a_k = 1,\;\; 0 < a_k \leq 1$$

We want to show that

$$\lim_{n\to\infty} E[\tilde{U}_n(y\hat{Y}^n)] = \lim_{n\to\infty} \tilde{v}_n(y)$$

We have

$$\begin{aligned}\tilde{v}_n(y) \leq E[\tilde{U}_n(y\hat{Y}^n)] = E[\tilde{U}_n(y\sum_{k\geq n} a_k Y_T^k)] &\leq \sum_{k\geq n} a_k E[\tilde{U}_n(yY_T^k)] \\ &\leq \sum_{k\geq n} a_k E[\tilde{U}_k(yY_T^k)] \\ &\leq \sup_{k\geq n} E[\tilde{U}_k(yY_T^k)],\end{aligned}$$

which hold true, since $\tilde{U}_n$ is an increasing and convex sequence.

On the one hand, for $k \geq n$, we have

$$E[\tilde{U}_k(yY_T^k)] \leq \tilde{v}_k(y) + \frac{1}{k} \leq (\sup_{k\geq n} \tilde{v}_k)(y) + \frac{1}{n}$$

which implies

$$\sup_{k\geq n} E[\tilde{U}_k(yY_T^k)] \leq (\sup_{k\geq n} \tilde{v}_k)(y) + \frac{1}{n} = (\lim_{n\to\infty} \tilde{v}_n)(y) + \frac{1}{n},$$

Therefore, we have

$$\tilde{v}_n(y) \leq E[\tilde{U}_n(y\hat{Y}^n)] \leq (\lim_{n\to\infty} \tilde{v}_n)(y) + \frac{1}{n},$$

by taking the limits, we obtain the result

$$\lim_{n\to\infty} \tilde{v}_n(y) \leq \lim_{n\to\infty} E[\tilde{U}_n(y\hat{Y}^n)] \leq (\lim_{n\to\infty} \tilde{v}_n)(y).$$

We have shown that $\lim_{n\to\infty} \tilde{v}_n(y) = \lim_{n\to\infty} E[\tilde{U}_n(y\hat{Y}^n)]$. Now we want to show that

$$\lim_{n\to\infty} \tilde{v}_n(y) = \tilde{v}(y).$$

We want to prove that $\tilde{U}_n(y) = \tilde{U}(y)$ for $y > U'(n)$ $(U'(n) \longrightarrow 0)$ as $n \to \infty$. U' is strictly decreasing from $(0, \infty)$ to $(0, \infty)$, $U'(n+1) \leq U'(n)$ and $U'(\infty) = 0$.

We want to show that

$$\tilde{U}_n(z + U'(n)) = \tilde{U}(z + U'(n)) \quad \forall z > 0.$$

We recall that

$$\tilde{U}(y) = \sup_{x>0}[u(x) - xy], \qquad \tilde{U}_n(y) = \sup_{0<x\le n}[u(x) - xy].$$

It is clear that

$$\tilde{U}(z + U'(n)) \ge \tilde{U}_n(z + U'(n))$$

Conversely,

$$\begin{aligned} \tilde{U}(z + U'(n)) = \sup_{x>0}[u(x) - x(z + U'(n))] &= \sup_{0<x\le I(z+U'(n))}[u(x) - x(z + U'(n))] \\ &\le \sup_{0<x\le n}[u(x) - x(z + U'(n))] \\ &\le \sup_{0<x\le m}[u(x) - x(z + U'(n))] \\ &= \tilde{U}_m(z + U'(n)), \end{aligned}$$

since we have I is decreasing and $z + U'(n) \ge U'(n)$, which implies

$$I(z + U'(n)) \le I(U'(n)) = n \le m.$$

We have by Fatou's Lemma and the uniformly integrability

$$\begin{aligned} \tilde{v}(y) &\le E[\tilde{U}(yY_T)] \\ &\le \liminf_{n\to\infty} E[\tilde{U}(yY_T + U'(n))] \\ &\le \liminf_{n\to\infty} \liminf_{m\ge n} E[\tilde{U}(y\hat{Y}^m + U'(n))], \qquad \hat{Y}^m \to Y_T, \\ &\le \liminf_{n\to\infty} \liminf_{m,n\ge n} E[\tilde{U}_m(y\hat{Y}^m + U'(n))], \quad \tilde{U}_m(z) = \tilde{U}(z), \text{ for } z \ge U'(n) \\ &\le \liminf_{n\to\infty} \liminf_{m,n\ge n} E[\tilde{U}_m(y\hat{Y}^m)], \qquad z \to \tilde{U}_m(z) \text{ is decreasing, } z = y\hat{Y}^m + U'(n) \\ &\le \lim_{n\to\infty} \tilde{v}_n(y) \\ &\le \tilde{v}(y). \end{aligned}$$

By Fatou's Lemma, we first deduce that

$$
\begin{aligned}
\liminf_{n\to\infty} E\{\tilde{U}_n^+(y\hat{Y}^n)\} &\geq E\{\liminf_{n\to\infty} \tilde{U}_n^+(y\hat{Y}^n)\} \\
&= E\{\tilde{U}^+(yY_T)\},
\end{aligned}
$$

and on the other hand, by Lemma 3.4.4 the family $\{\tilde{U}^-(yY_T), Y_T \in \mathcal{D}\}$ is uniformly integrable, and we have

$$
\begin{aligned}
\liminf_{n\to\infty} E\{\tilde{U}_n^-(y\hat{Y}^n)\} &= E\{\liminf_{n\to\infty} \tilde{U}_n^-(y\hat{Y}^n)\} \\
&= E\{\tilde{U}^-(yY_T)\}.
\end{aligned}
$$

Since $\tilde{U}_n$ is convex, we have

$$
\begin{aligned}
\lim_{n\to\infty} \tilde{v}_n(y) &= \lim_{n\to\infty} E\{\tilde{U}_n(yY^n)\} \\
&\geq \liminf_{n\to\infty} E\{\tilde{U}_n(y\hat{Y}^n)\} \\
&= \liminf_{n\to\infty} E\{\tilde{U}_n^+(y\hat{Y}^n)\} - \liminf_{n\to\infty} E\{\tilde{U}_n^-(y\hat{Y}^n)\} \\
&\geq E\{\liminf_{n\to\infty} \tilde{U}_n^+(y\hat{Y}^n)\} - E\{\liminf_{n\to\infty} \tilde{U}_n^-(y\hat{Y}^n)\} \\
&= E\{\tilde{U}^+(yY_T)\} - E\{\tilde{U}^-(yY_T)\} = E\{\tilde{U}(yY_T)\} \geq \tilde{v}(y),
\end{aligned}
$$

which prove that (3.80) holds true and therefore the proof of the second equation of system of (3.68). Under the assumption (3.1.2), the relation in the first equation in system of (3.68) is a consequence of the bipolarity or double duality property of the Fenchel-Legendre transform for convex functions (see Theorem B.2.2 for details). ∎

3.4.1 Study of Dual Problem

In this subsection we point out some properties of the dual objective function $\tilde{v}$ and establish the existence (and uniqueness) of $\hat{Y}_T^y \in \mathcal{D}$ which is optimal in the dual problem (3.42).

Proposition 3.4.9 *(Existence of a solution to the dual problem)*

Let U be a utility function satisfying (3.1.1), (3.1.2) and (3.3.1). Then, for all $y \in dom(\tilde{v})$ there exists a unique solution $\hat{Y}_T^y \in \mathcal{D}$ to $\tilde{v}(y)$. In particular, $\tilde{v}$ is strictly convex on $dom(\tilde{v})$.

Proof. For all $y \in dom(\tilde{v})$, let $(Y_T^n)_{n\geq 1}$ be a minimizing sequences in $\mathcal{D}$ of $\tilde{v}_n(y) < \infty$:

$$
\lim_{n\to\infty} \tilde{v}_n(y) = \tilde{v}(y)
$$

From the compactness theorem in $L^0_+(\Omega, \mathcal{F}_T, P)$, we can find a convex combination $\hat{Y}^n \in conv(Y^n_T, Y^{n+1}_T, \cdots)$, which is still lying in the convex set $\mathcal{D}$, and converges a.s. to a non-negative random variable Y^y_T . Since $\mathcal{D}$ is closed for the convergence in measure, we have $Y^y_T \in \mathcal{D}$. Analysis similar to that in the proof of Proposition 3.4 shows that

$$\begin{aligned}
\tilde{v}(y) = \lim_{n\to\infty} \tilde{v}_n(y) &\geq \liminf_{n\to\infty} E\{\tilde{U}_n(y\hat{Y}^n)\} \\
&\geq E\{\liminf_{n\to\infty} \tilde{U}_n(y\hat{Y}^n)\} \\
&\geq E\{\tilde{U}(y\hat{Y}^y_T)\} \geq \tilde{v}(y),
\end{aligned}$$

which prove that $\hat{Y}^y_T$ is a solution of $\tilde{v}(y)$. The uniqueness property is an immediate consequence from the strict convexity property of $\tilde{U}$ which is inherited by $\tilde{v}$ on its domain dom($\tilde{v}$). ■

The following lemma provides an easy and useful characterization of the reasonable asymptotic elasticity condition $AE(U) < 1$ in terms of conditions involving the functions U, V or the derivatives U', $V' = -I$ respectively.

Lemma 3.4.10 *Let U be a utility function satisfying (3.1.1) and (3.3.1). Then, the following assertions are equivalent:*

(i) $AE(U) < 1$

(ii) *There exist $x_0 > 0$ and $\gamma \in (0,1)$ such that*

$$xU^{'}(x) < \gamma U(x), \qquad \forall x \geq x_0.$$

(iii) *There exist $x_0 > 0$ and $\gamma \in (0,1)$ such that*

$$U(\lambda x) < \lambda^\gamma U(x), \qquad \forall x \geq x_0, \ \forall \lambda > 1.$$

(iv) *There exist $y_0 > 0$ and $\gamma \in (0,1)$ such that*

$$\widetilde{U}(\mu y) < \mu^{-\frac{\gamma}{1-\gamma}} \widetilde{U}(y), \ \forall 0 < \mu < 1, \qquad \forall 0 < y \leq y_0.$$

(v) *There exist $y_0 > 0$ and $\gamma \in (0,1)$ such that*

$$-y\widetilde{U}^{'}(y) < \frac{\gamma}{1-\gamma}\widetilde{U}(y), \ , \qquad \forall 0 < y \leq y_0.$$

Proof. Let us prove the equivalence $(i) \Leftrightarrow (ii)$. Suppose there is an $x_0 > 0$ and $\gamma \in (0,1)$ with the property that for all $x \geq x_0$ we have $xU^{'}(x) < \gamma U(x)$. The function U being strictly increasing with $U(\infty) > 0$, we may found a large $z > x_0$ such that $U(x) > 0, \forall\, x \geq z$, this insure that $U(x) \neq 0, \forall\, x \geq z$, therefore

$$\frac{xU^{'}(x)}{U(x)} < \gamma < 1, \qquad \forall x \geq z > x_0\,,$$

and we have

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} \leq \sup_{x\geq z} \frac{xU'(x)}{U(x)} < \gamma < 1\,. \tag{3.81}$$

For the converse, we take $\gamma_1 := AE(U) < 1$, and we can find an $\varepsilon > 0$ such that for $\gamma := \gamma_1 + \varepsilon < 1$, there exists $x_0 > 0$ such that for all $x \geq x_0$ we have $xU^{'}(x) < \gamma U(x)$.

$(ii) \Leftrightarrow (iii)$: For $x \geq x_0$ fixed, we define for $\lambda \in [1,\infty)$ the differentiable functions $F(\lambda) = U(\lambda x)$ and $G(\lambda) = \lambda^{\gamma} U(x)$. Notice that $F(1) = G(1)$ and that condition (iii) is equivalent to

$$F(\lambda) < G(\lambda), \qquad \forall \lambda \in [1,\infty),$$

for all $x \geq x_0$. Suppose that (ii) holds true. Then, $F'(1) < G'(1)$, and we conclude by the continuity of F' and G', that there is an $\varepsilon > 0$ with the property that for all $\lambda \in (1, 1+\varepsilon]$ we have $F'(\lambda) < G'(\lambda)$. Therefore, for all $\lambda \in (1, 1+\varepsilon]$, we have

$$F(\lambda) = F(1) + \int_1^{\lambda} F'(y)dy < G(\lambda) = G(1) + \int_1^{\lambda} G'(y)dy\,.$$

We now proceed to prove that this strict inequality holds for all $\lambda > 1$. Suppose this is not the case, we would have

$$\hat{\lambda} := \inf\{\lambda > 1 : F(\lambda) = G(\lambda)\} < \infty.$$

But at this point $\hat{\lambda}$ we should have $F'(\hat{\lambda}) \geq G'(\hat{\lambda})$. But from (ii)

$$F'(\hat{\lambda}) = xU'(\hat{\lambda}x) < \frac{\gamma}{\hat{\lambda}} U(\hat{\lambda}x) = \frac{\gamma}{\hat{\lambda}} F(\hat{\lambda}) = \frac{\gamma}{\hat{\lambda}} G(\hat{\lambda}) = G'(\hat{\lambda}),$$

which is the required contradiction. Conversely, suppose that (iii) is valid. Then, $F'(1) \leq G'(1)$ and we have

$$U'(x) = \frac{F'(1)}{x} \leq \frac{G'(1)}{x} = \gamma \frac{U(x)}{x},$$

which implies (ii).

$(iv) \Leftrightarrow (v)$: Let us fixing $0 < y \leq y_0$, by considering the associated differentiable functions $F(\mu) = \tilde{U}(\mu y)$ and $G(\mu) = \mu^{-\frac{\gamma}{1-\gamma}}\tilde{U}(y)$ $\forall \mu \in [0,1)$ with $F(1) = G(1)$. Notice that condition (iv) is equivalent to

$$F(\mu) < G(\mu), \qquad \forall \mu \in [0,1),$$

for all $0 < y \leq y_0$. Suppose that condition (v) holds true. Then we have $G'(1) < F'(1)$ and we deduced that there exist $\varsigma > 0$ such that for all $\mu \in [1-\varsigma, 1)$ $F(\mu) < G(\mu)$. We now prove that it remains valid for all $\mu < 1$. On the contrary, this would mean that

$$\hat{\mu} := \sup\{\mu < 1 : F(\mu) = G(\mu) < \infty\}.$$

At this point $\hat{\mu}$ we will have $F'(\hat{\mu}) \leq G'(\hat{\mu})$. But from (iv)

$$\begin{aligned} F'(\hat{\mu}) = y\tilde{U}'(\hat{\mu}y) &\geq y\frac{\hat{\mu}}{\mu}\tilde{U}'(\hat{\mu}y) \\ &\geq -\frac{\gamma}{1-\gamma}\tilde{U}'(\hat{\mu}y) \\ &\geq -\frac{\gamma}{1-\gamma}\hat{\mu}^{-\frac{\gamma}{1-\gamma}}\tilde{U}(y) \\ &\geq -\frac{\gamma}{1-\gamma}\hat{\mu}^{-\frac{1}{1-\gamma}}\tilde{U}(y) = G'(\hat{\mu}) \end{aligned}$$

which is the required contradiction. Conversely, suppose (iv); hold. then, $G'(1) < F'(1)$ and we have

$$\tilde{U}'(y) = \frac{F'(1)}{y} > \frac{G'(1)}{y} = -\frac{\gamma}{1-\gamma}\frac{\tilde{U}(y)}{y},$$

which means that

$$y\tilde{U}'(y) > -\frac{\gamma}{1-\gamma}\tilde{U}(y),$$

thus we obtain assertion (v).

$(ii) \Leftrightarrow (v)$: Suppose (ii), and put $y_0 = U'(x_0)$. Then, for all $0 < y < y_0$, we have $I(y) > I(y_0) = x_0$, since I is strictly decreasing, and so, we apply (ii) and we found that

$$\tilde{U}(y) = U(I(y)) - yI(y) > \frac{I(y)}{\gamma}U'(I(y)) - yI(y) = (\frac{1}{\gamma} - 1)yI(y) = (\frac{1-\gamma}{\gamma})yI(y),$$

since $\widetilde{U}'(I(y)) = y$, assertion (v) is proved. Conversely, suppose (v), and set $x_0 = I(y_0) =$

$-\widetilde{U}'(y_0)$. Then, for all $x \geq x_0$, we have $U'(x) \leq U'(x_0) = y_0$, and so

$$\begin{aligned} U(x) = \widetilde{U}(U'(x)) + xU'(x) &> -\frac{1-\gamma}{\gamma}U'(x)\widetilde{U}'(U'(x)) + xU'(x) \\ &= \frac{1-\gamma}{\gamma}U'(x)I(U'(x)) + xU'(x), \qquad \widetilde{U}' = -I \\ &= \frac{1-\gamma}{\gamma}xU'(x) + xU'(x), \qquad I = (U')^{-1} \\ &= \{\frac{1-\gamma}{\gamma} + 1\}xU'(x) \\ &= \frac{1}{\gamma}xU'(x), \end{aligned}$$

which is exactly assertion (ii). ■

Remark 3.4.11 *1. The characterizations* (iv) *and* (v) *indicate that if* $AE(U) < 1$*, then there is* $y_0 > 0$ *with the property that for every* $0 < \mu < 1$ *we have*

$$yI(\mu y) \leq C\widetilde{U}(y), \quad 0 < y \leq y_0, \tag{3.82}$$

where the positive constant C *depends on* μ.

2. Remark that the characterization (ii) *for the reasonable asymptotic elasticity condition* $AE(U) < 1$ *yields the growth condition (3.36) on the utility function* U *as indicated in Remark 3.3.3.*

The following proposition states the main properties of the value function $\tilde{v}$ and provides a probabilistic characterization of the solution $\hat{Y}_T^y$ to the dual problem (3.42) by deriving the first-order (i.e. on the first derivative of $\tilde{v}$) optimality conditions.

Proposition 3.4.12 *(Characterization of the solution to the dual problem)*

Let U *be a utility function satisfying (3.1.1), (3.1.2), (3.3.1) and (3.3.7). Then, the value function* $\tilde{v}$ *of the dual problem is finitely valued, continuously differentiable and strictly convex on* $(0,\infty)$ *with*

$$\begin{aligned} -\tilde{v}'(y) &= E[\hat{Y}_T^y I(y\hat{Y}_T^y)] \\ &= \sup_{Y_T \in \mathcal{D}} E[Y_T I(y\hat{Y}_T^y)], \qquad y > 0, \end{aligned} \tag{3.83}$$

and thus,

$$I(y\hat{Y}_T^y) \in \mathcal{C}(-\tilde{v}'(y)), \qquad y > 0. \tag{3.84}$$

Proof. We first prove that $\tilde{v}$ is finite and $dom(\tilde{v}(y)) = (0, \infty)$.

From assumption in (3.1.2) and the conjugate duality (3.68) it follow that for $x > 0$ such that $v(x) < \infty$, we have $\inf_{y>0}\{\tilde{v}(y) + yx\} < \infty$. Therefore, there exists a $y_1 > 0$ such that $\tilde{v}(y_1) < \infty$, i.e. $dom(\tilde{v}(y)) \neq \emptyset$.

Note that the decreasing property of $\widetilde{U}$ is inherited by $\tilde{v}$, therefore

$$\tilde{v}(y) \leq \tilde{v}(y_1) < \infty\,, \quad \text{for all} \quad y \geq y_1\,.$$

Let us now prove that for $y \in (0, y_1)$, we also have $\tilde{v}(y) < \infty$. The fact that $\tilde{v}(y_1) < \infty$ implies there exists $Y_T \in \mathcal{D}$ such that $E[\tilde{U}(y_1 Y_T)] < \infty$. From (3.68) we also have for a given $x_0 > 0$

$$\tilde{U}(y_1 Y_T) \geq U(x_0) - x_0 y_1 Y_T, \qquad \text{with} \quad E[Y_T] \leq 1,$$

which prove that $\tilde{U}(y_1 Y_T) \in L^1(P)$.

We have

$$\tilde{U}(yY_T) = \tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T \leq y_0} + \tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T > y_0}$$

Firstly, we deal with $\tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T > y_0}$. On the event $\{y_1 Y_T > y_0\} = \{yY_T > \dfrac{yy_0}{y_1}\}$, we have

$$\tilde{U}(yY_T) \leq \tilde{U}(\frac{yy_0}{y_1}), \quad \text{since } \tilde{U} \text{ is decreasing and } 0 < y < y_1$$

which implies

$$\tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T > y_0} \leq \tilde{U}(\frac{yy_0}{y_1})\mathbf{1}_{y_1 Y_T > y_0} \leq |\tilde{U}(\frac{yy_0}{y_1})|$$

Now let us consider the last part $\tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T \leq y_0}$. The characterisation (iv) of $AE(U) < 1$ in Lemma 3.4.10 indicates that there is a $y_0 > 0$ such that $\tilde{U}(z\mu) \leq C(\mu)\tilde{U}(z)$ for all $0 < \mu < 1$ and all $0 < z \leq y_0$. This inequality clearly holds when $z = 0$. Since we are considering the case $0 < y < y_1$ and we are working on the event $\{y_1 Y_T \leq y_0\}$ we choose $0 < \mu = \frac{y}{y_1} < 1$ and $z = y_1 Y_T \leq y_0$, it follows that there is a $y_0 > 0$ with the property that for all $0 < y < y_1$, we have

$$\tilde{U}(yY_T) \leq C(\frac{y}{y_1})\tilde{U}(y_1 Y_T),$$

which implies

$$\begin{aligned}\tilde{U}(yY_T)\mathbf{1}_{y_1 Y_T \leq y_0} &\leq C(y)\tilde{U}(y_1 Y_T)\mathbf{1}_{y_1 Y_T \leq y_0} \\ &\leq C(y)|\tilde{U}(y_1 Y_T)|\end{aligned}$$

By combining the above two inequalities, it yields

$$\begin{aligned}\tilde{U}(yY_T) &\leq C(y)\tilde{U}(y_1Y_T)\mathbf{1}_{y_1Y_T\leq y_0} + \tilde{U}(yY_T)\mathbf{1}_{y_1Y_T>y_0}\\ &\leq C(y)|\tilde{U}(y_1Y_T)| + |\tilde{U}(\frac{y}{y_1}y_0)|,\end{aligned}$$

for some positive constant $C(y)$. Therefore,

$$\tilde{v}(y) \leq E[\widetilde{U}(yY_T)] \leq C(y)E[|\widetilde{U}(y_1Y_T)|] + |\widetilde{U}(\frac{y}{y_1}y_0)| < \infty,$$

Thus, we just prove that $\tilde{v}(y) < \infty$ for $y < y_1$ and therefore $\tilde{v}(y) < \infty$ for $y > 0$, which prove that $dom(\tilde{v}(y)) = (0, \infty)$.

The function $\widetilde{U}$ is decreasing and strictly convex. These properties in conjunction with Jensen's inequality yield

$$E(\widetilde{U}(yY_T)) \geq \widetilde{U}(yE(Y_T)) \geq \widetilde{U}(xy) > -\infty \tag{3.85}$$

hence $\tilde{v}(y) \geq \widetilde{U}(xy) > -\infty$.

Now, we want to prove that $\tilde{v}$ is continuously differentiable. It is enough, by convexity of $\tilde{v}$, to prove the existence of its derivative everywhere on $(0, \infty)$. Let us fix $y > 0$. Then, for all $\alpha > 0$, using the definition of $\tilde{v}$ and the convexity of $\widetilde{U}$, we have

$$\begin{aligned}\frac{\tilde{v}(y+\alpha) - \tilde{v}(y)}{\alpha} &= \inf_{Y_T\in\mathcal{D}} E[\frac{\widetilde{U}((y+\alpha)Y_T)}{\alpha}] - E[\frac{\widetilde{U}(\hat{Y}_T^y)}{\alpha}]\\ &\leq E[\frac{\widetilde{U}((y+\alpha)\hat{Y}_T^y) - \tilde{U}(y\hat{Y}_T^y)}{\alpha}]\\ &\leq E[\hat{Y}_T^y\widetilde{U}'((y+\alpha)\hat{Y}_T^y)].\end{aligned}$$

By using the fact that $-\tilde{U}' = I$ and Fatou's Lemma

$$\begin{aligned}\limsup_{\alpha\downarrow 0}\frac{\tilde{v}(y+\alpha) - \tilde{v}(y)}{\alpha} &\leq \limsup_{\alpha\downarrow 0} E[\hat{Y}_T^y\tilde{U}'((y+\alpha)\hat{Y}_T^y)]\\ &\leq -E[\hat{Y}_T^y I(\limsup_{\alpha\downarrow 0}(y+\alpha)\hat{Y}_T^y)]\\ &= -E[\hat{Y}_T^y I(y\hat{Y}_T^y)].\end{aligned}\tag{3.86}$$

By the same argument as above, for $\alpha > 0$ such that $y - \alpha > 0$ we have

$$\begin{aligned}\frac{\tilde{v}(y) - \tilde{v}(y-\alpha)}{\alpha} &= \inf_{Y_T\in\mathcal{D}} E[\frac{\widetilde{U}(y\hat{Y}_T^y) - \widetilde{U}((y-\alpha)Y_T)}{\alpha}]\\ &\geq E[\frac{\tilde{U}(y\hat{Y}_T^y) - \tilde{U}((y-\alpha)\hat{Y}_T^y)}{\alpha}]\\ &\geq E[\hat{Y}_T^y\tilde{U}'((y-\alpha)\hat{Y}_T^y)].\end{aligned}\tag{3.87}$$

The characterizations (iv) and (v) in Lemma 3.4.10 show that if $AE(U) < 1$, then there is a $y_0 > 0$ with the property that for all $0 < \mu < 1$

$$yI(\mu y) \le C\tilde{U}(y), \qquad 0 \le y \le y_0,$$

where the positive constant C depends on μ.

We know that $E[\tilde{U}(yY_T)] < \infty$ and thus $\tilde{U}(yY_T) \in L^1(P)$. Using the above result there exist $y_0 > 0$ with the property that for all $0 < \delta < y/2$ i.e. $\frac{1}{2} < 1 - \delta/y < 1$ We have

$$\hat{Y}^y_T I((y-\delta)\hat{Y}^y_T) = \hat{Y}^y_T I((y-\delta)\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T \le y_0} + \hat{Y}^y_T I((y-\delta)\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T > y_0}$$

Let us considering the event $\{y\hat{Y}^y_T > y_0\}$.

$$\begin{aligned}
y\hat{Y}^y_T > y_0 &\implies \hat{Y}^y_T > \frac{y_0}{y} \\
&\implies (y-\delta)\hat{Y}^y_T > (y-\delta)\frac{y_0}{y} = y_0(1-\frac{\delta}{y}) \ge \frac{y_0}{2} \\
&\implies I(\frac{y_0}{2}) \ge I((y-\delta)\hat{Y}^y_T), \quad \text{since I is decreasing} \\
&\implies \hat{Y}^y_T I((y-\delta)\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T > y_0} \le \hat{Y}^y_T I(\frac{y_0}{2})\mathbf{1}_{y\hat{Y}^y_T > y_0} \le \hat{Y}^y_T I(\frac{y_0}{2})
\end{aligned}$$

Now, let us consider the event $\{y\hat{Y}^y_T \le y_0\}$. There exist $y_0 > 0$ such that for all $0 < \mu < 1$

$$zI(\mu z) \le C(\mu)\tilde{U}(z) \quad \forall 0 < z \le y_0.$$

We take $z = y\hat{Y}^y_T$ and $0 < \mu = 1 - \dfrac{\delta}{y} < 1$ and there exists $y_0 > 0$ with the property that

$$y\hat{Y}^y_T I((1-\frac{\delta}{y})\hat{Y}^y_T) \le C(y)\widetilde{U}(y\hat{Y}^y_T).$$

we have

$$\begin{aligned}
\hat{Y}^y_T I((y-\delta)\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T \le y_0} &= \frac{1}{y}\big[y\hat{Y}^y_T I((1-\frac{\delta}{y})\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T \le y_0}\big] \\
&\le \frac{1}{y}C(y)\widetilde{U}(y\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T \le y_0} \\
&\le C(y)|\widetilde{U}(y\hat{Y}^y_T)|.
\end{aligned}$$

Considering the two above result, it follow that

$$\begin{aligned}
0 &\le -\hat{Y}^y_T\widetilde{U}'((y-\delta)\hat{Y}^y_T) = \hat{Y}^y_T I((y-\delta)\hat{Y}^y_T) \\
&\le C(y)\widetilde{U}(y\hat{Y}^y_T)\mathbf{1}_{y\hat{Y}^y_T \le y_0} + \hat{Y}^y_T I(\frac{y_0}{2})\mathbf{1}_{y\hat{Y}^y_T > y_0} \\
&\le C(y)|\widetilde{U}(y\hat{Y}^y_T)| + \hat{Y}^y_T I(\frac{y_0}{2}),
\end{aligned}$$

with $C(y) < \infty$ and y_0 arbitrary.

The right hand side of this last inequality is an integrable random variable, and therefore Lebesgue dominated convergence theorem may be applied to inequality (3.87) by letting α tend to zero:

$$\begin{aligned}\liminf_{\alpha\downarrow 0}\frac{\tilde{v}(y)-\tilde{v}(y-\alpha)}{\alpha} &\geq -\liminf_{\alpha\downarrow 0} E[\hat{Y}_T^y\tilde{U}'((y-\alpha)\hat{Y}_T^y)] \\ &\geq -E[\hat{Y}_T^y I(\liminf_{\alpha\downarrow 0}(y-\alpha)\hat{Y}_T^y)] \\ &= -E[\hat{Y}_T^y I(y\hat{Y}_T^y)].\end{aligned} \tag{3.88}$$

From the convexity of $\tilde{v}$ we obtain the existence of the one side derivative

$$\tilde{v}'(y-) \leq \tilde{v}'(y+)$$

From (3.86) we obtain

$$\limsup_{\delta\downarrow 0}\frac{\tilde{v}(y+\delta)-\tilde{v}(y)}{\delta} \leq -E[\hat{Y}_T^y I(y\hat{Y}_T^y)]$$

which implies

$$\tilde{v}'(y+) \leq -E[\hat{Y}_T^y I(y\hat{Y}_T^y)].$$

From (3.88) we obtain

$$\liminf_{\delta\downarrow 0}\frac{\tilde{v}(y)-\tilde{v}(y-\delta)}{\delta} \geq -E[\hat{Y}_T^y I(y\hat{Y}_T^y)],$$

which implies

$$\tilde{v}'(y-) \geq -E[\hat{Y}_T^y I(y\hat{Y}_T^y)] \geq \tilde{v}'(y+).$$

Thus, from (3.86), (3.88) and the convexity of $\tilde{v}$ we conclude that $\tilde{v}$ is differentiable at any $y \in (0,\infty)$ and with

$$\tilde{v}'(y) = -E[\hat{Y}_T^y I(y\hat{Y}_T^y)]. \tag{3.89}$$

We want to prove that

$$E[\hat{Y}_T^y I(y\hat{Y}_T^y)] = \sup_{Y_T\in\mathcal{D}} E[Y_T I(y\hat{Y}_T^y)]. \tag{3.90}$$

It is clear that $\hat{Y}_T^y \in \mathcal{D}$ and

$$E[\hat{Y}_T^y I(y\hat{Y}_T^y)] \leq \sup_{Y_T \in \mathcal{D}} E[Y_T I(y\hat{Y}_T^y)]. \tag{3.91}$$

Conversely, we consider an arbitrary element $Y_T \in \mathcal{D}$ and we define

$$Y_T^\varepsilon = (1-\varepsilon)\hat{Y}_T^y + \varepsilon Y_T \in \mathcal{D}, \;\; 0 < \varepsilon < 1 \tag{3.92}$$

Notice that $\lim_{\varepsilon \downarrow 0} Y_T^\varepsilon = \hat{Y}_T^y$. By definition of $\tilde{v}$

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E[\tilde{U}(yY_T)] = E[\tilde{U}(y\hat{Y}_T^y)] \leq E[\tilde{U}(y\hat{Y}_T^\varepsilon)]. \tag{3.93}$$

Using (3.92) and the decreasing property of $\widetilde{U}'$, we have

$$Y_T^\varepsilon = (1-\varepsilon)\hat{Y}_T^y + \varepsilon Y_T \geq (1-\varepsilon)\hat{Y}_T^y, \tag{3.94}$$

implies

$$\widetilde{U}'(Y_T^\varepsilon) \leq \widetilde{U}'((1-\varepsilon)\hat{Y}_T^y). \tag{3.95}$$

For any convex function $\widetilde{U}$, we have

$$\widetilde{U}(x) \geq \widetilde{U}(y) + (x-y)\widetilde{U}'(y)$$

Therefore, we also have

$$\widetilde{U}(y\hat{Y}_T^y) - \widetilde{U}(yY_T^\varepsilon) = \int_{yY_T^\varepsilon}^{y\hat{Y}_T^y} \widetilde{U}'(z)dz \geq \tilde{U}'(yY_T^\varepsilon)[y\hat{Y}_T^y - yY_T^\varepsilon] \tag{3.96}$$

By using (3.93) and (3.96), we have

$$E[\tilde{U}'(yY_T^\varepsilon)y(\hat{Y}_T^y - Y_T^\varepsilon)] \leq 0, \tag{3.97}$$

which implies

$$E[I(yY_T^\varepsilon)y(\hat{Y}_T^y - Y_T^\varepsilon)] \geq 0, \tag{3.98}$$

where $\tilde{U}' = -I$. On the one hand, since $\varepsilon(\hat{Y}_T^y - Y_T) = \hat{Y}_T^y - Y_T^\varepsilon$, we have

$$E[I(yY_T^\varepsilon)(\hat{Y}_T^y - Y_T)] \geq 0, \tag{3.99}$$

which implies,

$$E[I(yY_T^\varepsilon)\hat{Y}_T^y] \geq E[I(yY_T^\varepsilon)Y_T)]. \tag{3.100}$$

On the other hand, we know that,

$$yY_T^\varepsilon \geq y(1-\varepsilon)\hat{Y}_T^y \tag{3.101}$$

which implies

$$I(yY_T^\varepsilon) \leq I(y(1-\varepsilon)\hat{Y}_T^y), \tag{3.102}$$

since I is decreasing. By (3.100) and (3.102) we have

$$E[\hat{Y}_T^y I(y(1-\varepsilon)\hat{Y}_T^y)] \geq E[Y_T I(yY_T^\varepsilon)]. \tag{3.103}$$

By Fatou Lemma, we have

$$\begin{aligned} \liminf_{\varepsilon \downarrow 0} E[Y_T I(y\hat{Y}_T^\varepsilon)] &\geq E[\liminf_{\varepsilon \downarrow 0} Y_T I(y\hat{Y}_T^\varepsilon)] \\ &= E[Y_T I(y\hat{Y}_T^y)]. \end{aligned}$$

By the convergence dominated theorem, we have

$$E[\hat{Y}_T^y I(y(1-\varepsilon)\hat{Y}_T^y)] \longrightarrow E[\hat{Y}_T^y I(y\hat{Y}_T^y)], \tag{3.104}$$

which implies

$$E[\hat{Y}_T^y I(y\hat{Y}_T^y)] \geq E[Y_T I(y\hat{Y}_T^y)], \tag{3.105}$$

where Y_T is arbitrary, and we obtain the result

$$E[\hat{Y}_T^y I(y\hat{Y}_T^y)] \geq \sup_{Y_T \in \mathcal{D}} E[Y_T I(y\hat{Y}_T^y)]. \tag{3.106}$$

■

3.4.2 Proof of the Theorem 3.4.1

Proof. *(1)* Using the first equation (3.68) and the fact that $\tilde{v}$ is strictly convex on $(0,\infty)$ we see that v is differentiable on the open interval $(0,\infty)$. Moreover, the strict concavity of v on the $(0,\infty)$ is due to the fact that U is strictly concave and to the uniqueness of a solution to $v(x)$ as detailed below. In fact, using the definition of $v(x)$ in 3.2, we obtain easily the concavity of v. Indeed, for $x_1 > 0$, $x_2 > 0$ and for $\varepsilon \in (0,1)$, we have

$$\begin{aligned} v(\varepsilon x_1 + (1-\varepsilon)x_2) &= \sup_{\alpha\in\mathcal{A}(S)} E[U(\varepsilon x_1 + (1-\varepsilon)x_2 + \int_0^T \alpha_s dS_s)], \\ &\geq \varepsilon \sup_{\alpha\in\mathcal{A}(S)} E[U(x_1 + \int_0^T \alpha_s dS_s)] \\ &+ (1-\varepsilon) \sup_{\alpha\in\mathcal{A}(S)} E[U(x_2 + \int_0^T \alpha_s dS_s)] \\ &= \varepsilon v(x_1) + (1-\varepsilon)v(x_2), \end{aligned}$$

which is true since U is concave and therefore, v is is concave as stated.

We want to prove that v is strictly increasing and strictly concave. Let us first prove the strict increasing feature, that is for $0 < x_1 < x_2$, we have $v(x_1) < v(x_2)$. In fact, for $0 < x_1 < x_2$, we have $\frac{x_2}{x_1} > 1$, therefore $\hat{X}_T^{x_1} < \hat{X}_T^{x_1}\frac{x_2}{x_1}$ since $\hat{X}_T^{x_1} \geq 0$. The utility function U is strictly increasing by assumption, hence

$$v(x_1) = E[U(\hat{X}_T^{x_1})] < E[U(\hat{X}_T^{x_1}\frac{x_2}{x_1})] \leq v(x_2),$$

The second inequality holds true since $\hat{X}_T^{x_1} \in \mathcal{C}(x_1)$ and by consequence

$$E[\hat{X}_T^{x_1}\frac{x_2}{x_1}Y_T] = \frac{x_2}{x_1}E[\hat{X}_T^{x_1}Y_T] \leq \frac{x_2}{x_1}x_1 = x_2, \qquad \forall\, Y_T \in \mathcal{D}.$$

therefore $\hat{X}_T^{x_1}\frac{x_2}{x_1} \in \mathcal{C}(x_2)$. Thus, v is strictly increasing.

Now we want to show that v is strictly concave. Let us fix $0 < \varepsilon < 1$, $x_1, x_2 > 0$. We have already shown above that v is concave. We give another proof here based on the static characterisation of $v(x)$ as a supremum on the set $\mathcal{C}(x)$. For $i = 1,2$, we have shown that $\hat{X}_T^{x_i} \in \mathcal{C}(x_i)$ and $v(x_i) = E[U(\hat{X}_T^{x_i})]$. It is easy to check that

$$\frac{\hat{X}_T^{x_1} + \hat{X}_T^{x_2}}{2} \in \mathcal{C}(\frac{x_1+x_2}{2})\,.$$

Using the concavity of U, it follows that

$$\begin{aligned} v(\frac{x_1+x_2}{2}) &\geq E\big[U(\frac{\hat{X}_T^{x_1}+\hat{X}_T^{x_2}}{2})\big] \\ &\geq \frac{1}{2}E[U(\hat{X}_T^{x_1})]+\frac{1}{2}E[U(\hat{X}_T^{x_2})] \\ &= \frac{1}{2}v(x_1)+\frac{1}{2}v(x_2) \end{aligned}$$

Hence, v is a concave function, i.e. for $0<\varepsilon<1$ and $x_1, x_2>0$ we have

$$v(\varepsilon x_1+(1-\varepsilon)x_2) \geq \varepsilon v(x_1)+(1-\varepsilon)v(x_2)$$

To show the strict concavity of v, i.e. for $0<\varepsilon<1$ and $0<x_1<x_2$ we have

$$v(\varepsilon x_1+(1-\varepsilon)x_2) > \varepsilon v(x_1)+(1-\varepsilon)v(x_2)$$

we proceed by contradiction. Let us suppose there exist $0<x_1<x_2$ such that

$$v(\varepsilon x_1+(1-\varepsilon)x_2) = \varepsilon v(x_1)+(1-\varepsilon)v(x_2)\,,$$

holds. We have

$$\begin{aligned} v(\varepsilon x_1+(1-\varepsilon)x_2) &= E[U(\hat{X}_T^{\varepsilon x_1+(1-\varepsilon)x_2})] \\ &= \varepsilon E[U(\hat{X}_T^{x_1})]+(1-\varepsilon)E[U(\hat{X}_T^{x_2})]. \end{aligned}$$

We know that

$$\varepsilon\hat{X}_T^{x_1}+(1-\varepsilon)\hat{X}_T^{x_2} \in \mathcal{C}(\varepsilon x_1+(1-\varepsilon)x_2) \qquad \text{and } U \text{ is concave,}$$

which implies

$$U(\varepsilon\hat{X}_T^{x_1}+(1-\varepsilon)\hat{X}_T^{x_2}) \geq \varepsilon U(\hat{X}_T^{x_1})+(1-\varepsilon)U(\hat{X}_T^{x_2})$$

by consequence

$$v(\varepsilon x_1+(1-\varepsilon)x_2) \geq E[U(\varepsilon\hat{X}_T^{x_1}+(1-\varepsilon)\hat{X}_T^{x_2})] \geq \varepsilon v(x_1)+(1-\varepsilon)v(x_2)$$

Let us define

$$\begin{aligned} Z_1 &:= U(\varepsilon\hat{X}_T^{x_1}+(1-\varepsilon)\hat{X}_T^{x_2}) \\ Z_2 &:= \varepsilon U(\hat{X}_T^{x_1})+(1-\varepsilon)U(\hat{X}_T^{x_2}). \end{aligned}$$

we have shown that

$$E[Z_1 - Z_2] = 0 \text{ and } Z_1 - Z_2 \geq 0\,,$$

Therefore, $Z_1 = Z_2$ *a.s.*. The utility function U being strictly concave, it follows that $\hat{X}_T^{x_1} = \hat{X}_T^{x_2}$ *a.s.*, which implies that $v(x_1) = v(x_2)$, which is impossible since v is strictly increasing. Consequently, v is strictly concave.

The existence of the solution to $v(x)$ comes from the fact that $dom(\tilde{v}) = (0, \infty)$, which ensure assumption in 3.10. In fact, if we suppose that

$$\tilde{v}(y) < \infty \qquad \forall y > 0, \tag{3.107}$$

then, 3.70 implies that

$$\begin{aligned} \limsup_{x \longrightarrow \infty} \frac{v(x)}{x} &\leq \limsup_{x \longrightarrow \infty} \inf_{y>0} \{ \frac{\tilde{v}(y)}{x} + y \} \\ &\leq \inf_{y>0} \{ \limsup_{x \longrightarrow \infty} \frac{\tilde{v}(y)}{x} + y \} \\ &= \inf_{y>0} \{ y \} = 0, \end{aligned}$$

which in fact show that the assumption in 3.10 holds true. The condition 3.107 hold true if

$$\forall y > 0 \quad \exists Z_T \in \mathcal{M}_e : \ \ E[\tilde{U}(yZ_T)] < \infty. \tag{3.108}$$

(2) The existence and the uniqueness of the solution to $\tilde{v}(y)$ come from Proposition 3.4.9 and the fact that $dom(\tilde{v}) = (0, \infty)$.

(3) First of all, we need to check the condition

$$\tilde{v}'(\infty) := \lim_{y \longrightarrow \infty} \tilde{v}'(y) = 0.$$

The function $-\tilde{U}$ is increasing from $(0, \infty)$ into $(\tilde{U}(0), \infty)$ and the fact that $I(y) = -\tilde{U}'(y)$ implies that $-\tilde{U}$ converges to zero as y tends to infinity, since $I(\infty) = 0$. Thus, we have for all $\varepsilon > 0$ there exists $N_\varepsilon > 0$ such that

$$-\tilde{U}(y) \leq N_\varepsilon + \varepsilon y, \qquad y > 0.$$

By using the l'Hôpital rule, we found that

$$
\begin{aligned}
0 \leq \tilde{v}'(\infty) = \lim_{y \longrightarrow \infty} \frac{-\tilde{v}(y)}{y} &= \lim_{y \longrightarrow \infty} - \inf_{Y_T \in \mathcal{D}} E[\frac{\tilde{U}(yY_T)}{y}] \\
&= \lim_{y \longrightarrow \infty} \sup_{Y_T \in \mathcal{D}} E[\frac{-\tilde{U}(yY_T)}{y}] \\
&\leq \lim_{y \longrightarrow \infty} \sup_{Y_T \in \mathcal{D}} E[\frac{N_\varepsilon}{y} + \varepsilon Y_T] \\
&\leq \lim_{y \longrightarrow \infty} E[\frac{N_\varepsilon}{y} + \varepsilon] = \varepsilon,
\end{aligned}
$$

since, from 3.39 $E[Y_T] \leq 1$ for all $Y_T \in \mathcal{D}$ with $X_T = 1 \in \mathcal{C}(1)$. It thus follow that

$$\tilde{v}'(\infty) = 0. \tag{3.109}$$

Moreover, by the Proposition 3.4.12, we have

$$- \tilde{v}'(y) \geq E[Z_T I(y\hat{Y}_T^y)], \qquad \forall y > 0. \tag{3.110}$$

where $Z_T > 0$ a.s. is a fixed element in $\mathcal{M}_e$.

Remark that since $E[\hat{Y}_T^y] < 1$ for all $y > 0$, we have by Fatou Lemma $E[\hat{Y}_T^0] \leq 1$ where $\hat{Y}_T^0 = \liminf_{y \downarrow 0} \hat{Y}_T^y$. In particular, $\hat{Y}_T^0 < \infty$ a.s. By letting y tend to zero in 3.110, and recalling that $I(0) = \infty$, we obtain by Fatou's Lemma that

$$\tilde{v}'(0) = \lim_{y \downarrow 0} \tilde{v}'(y) = -\infty, \tag{3.111}$$

Using (3.109) and (3.111), we show now that for any $x > 0$ the strictly convex function $y \in (0, \infty) \longmapsto f_x(y) = \tilde{v}(y) + xy$ admits a unique minimum $\hat{y}$ such that $-\tilde{v}'(\hat{y}) = x$, which means that $\hat{y} = -(\tilde{v}')^{-1}(x) = v'(x)$ from the conjugate duality relation (3.69). Indeed, $\lim_{y \downarrow 0} f_x(y) = \tilde{v}(0+) = \widetilde{U}(0) = U(\infty)$. If $U(\infty) = \infty$, the $\inf f_x$ cannot be attained at $y = 0$. Suppose $U(\infty) < \infty$ and $\inf f_x$ is attained at $\hat{y} = 0$, i.e. $\inf_{y>0} f_x(y) = \inf_{y>0}(\tilde{v}(y) + xy) = \tilde{v}(0) = \widetilde{U}(0)$, this means

$$\tilde{v}(0) \quad \leq \quad \tilde{v}(y) + xy\,, \quad \forall y > 0 \tag{3.112}$$

This implies that

$$
\begin{aligned}
x \quad &\geq \quad \frac{\tilde{v}(0) - \tilde{v}(y)}{y}\,, \quad \forall y > 0 \ \forall Y_T \in \mathcal{D} &(3.113)\\
&\geq \quad -\tilde{v}'(0)\,, &(3.114)
\end{aligned}
$$

as y tends to zero, we obtain $x \geq \infty$, a clear contradiction.

Therefore, either the infimum of f_x is attained at a (unique) number $\hat{y} = \hat{y}_x \in (0, \infty)$ or it is attained at $\hat{y} = \infty$. If $\hat{y} = \infty$, there is a sequence $y_n \to \infty$ such that for a given and fixed $y \in (0, \infty)$, we have $y < y_n$ for large enough n, and therefore

$$f_x(y) = \tilde{v}(y) + xy \quad \geq \quad f_x(y_n) = \tilde{v}(y_n) + xy_n\,, \tag{3.115}$$

so

$$\begin{aligned} x &\leq \frac{\tilde{v}(y) - \tilde{v}(y_n)}{y_n - y}\,, && (3.116) \\ x &\leq -\tilde{v}'(\infty) = 0, && (3.117) \end{aligned}$$

where the last inequality follows by l'Hôpital rule when $y_n \to \infty$, therefore we have a contradiction.

Now, let us show that $I(\hat{y}\hat{Y}_T)$ is a solution to $v(x)$, i.e $I(\hat{y}\hat{Y}_T) = \hat{X}^x_T$. From Proposition 3.4.12, we have

$$I(\hat{y}\hat{Y}_T) \in \mathcal{C}(x) \quad \text{and} \quad E[\hat{Y}_T I(\hat{y}\hat{Y}_T)] = x.$$

It thus follow that from first 3.68

$$\begin{aligned} v(x) \geq E[U(I(\hat{y}\hat{Y}_T))] &= E[\widetilde{U}(\hat{y}\hat{Y}_T)] + E[\hat{y}\hat{Y}_T I(\hat{y}\hat{Y}_T)] \\ &= E[\widetilde{U}(\hat{y}\hat{Y}_T)] + x\hat{y} \\ &\geq \tilde{v}(\hat{y}) + x\hat{y}. \end{aligned}$$

By the conjugate relation 3.68 we have

$$v(x) \geq E[U(I(\hat{y}\hat{Y}_T))] \geq \tilde{v}(\hat{y}) + x\hat{y} = v(x),$$

and it follow that $v(x) = E[U(I(\hat{y}\hat{Y}_T))]$, which prove that $I(\hat{y}\hat{Y}_T)$ is a solution to $v(x)$.

(4) Let us suppose the existence of $y > 0$ such that $\inf_{Z_T \in \mathcal{M}_e} E[\widetilde{U}(yZ_T)] < \infty$, then we can find an element $Z^0_T \in \mathcal{M}_e$ such that $\widetilde{U}(yZ^0_T) \in L^1(P)$.

Let us consider $Y_T \in \mathcal{D}$, and let the sequence $(Z^n)_{n \geq 1}$ belongs to the set $\mathcal{M}_e$ and satisfies $Y_T \leq \lim_{n \to \infty} Z^n$ a.s. For all $\varepsilon \in (0, 1)$ and $n \geq 1$, we define the convex combination

$$\bar{Z}^{n,\varepsilon} = (1 - \varepsilon) Z^n + \varepsilon Z^0_T \, \in \mathcal{M}_e,$$

since $\mathcal{M}_e$ is a convex set. By multiplying this above equation by $y > 0$ we have

$$y\bar{Z}^{n,\varepsilon} = (1-\varepsilon)yZ^n + \varepsilon y Z_T^0 > \varepsilon y Z_T^0,$$

which implies by the decreasing property of $\widetilde{U}$ that

$$\tilde{U}(y\bar{Z}^{n,\varepsilon}) \leq \widetilde{U}(\varepsilon y Z_T^0) \tag{3.118}$$

In the one hand, using the characterization (iv) of $AE(U) < 1$ stated in Lemma 3.4.10, there exists $y_0 > 0$ with the property that on the event $\{yZ_T^0 \leq y_0\}$, we have

$$\widetilde{U}(\varepsilon y Z_T^0) \leq C_\varepsilon \widetilde{U}(yZ_T^0), \tag{3.119}$$

for some positive constant $C_\varepsilon > 0$.

In the other hand, by the decreasing property of $\widetilde{U}$, on the event $\{yZ_T^0 > y_0\}$, we have

$$\widetilde{U}(\varepsilon y_0) \geq \widetilde{U}(\varepsilon y Z_T^0). \tag{3.120}$$

Since

$$\tilde{U}(y\bar{Z}^{n,\varepsilon}) \leq \widetilde{U}(\varepsilon y Z_T^0) = \widetilde{U}(\varepsilon y Z_T^0)\mathbf{1}_{yZ_T^0 \leq y_0} + \widetilde{U}(\varepsilon y Z_T^0)\mathbf{1}_{yZ_T^0 > y_0}. \tag{3.121}$$

Replacing (3.120) and (3.119) into (3.121) we get

$$\widetilde{U}(y\bar{Z}^{n,\varepsilon}) \leq C_\varepsilon \tilde{U}(yZ_T^0)\mathbf{1}_{yZ_T^0 \leq y_0} + \tilde{U}(\varepsilon y_0)\mathbf{1}_{yZ_T^0 > y_0}. \tag{3.122}$$

We then get

$$\widetilde{U}^+(y\bar{Z}^{n,\varepsilon}) \leq C_\varepsilon |\tilde{U}(yZ_T^0)| + |\widetilde{U}(\varepsilon y_0)|, \qquad \forall n \geq 1, \tag{3.123}$$

which prove that the sequence of random variables $\{\widetilde{U}^+(y\bar{Z}^{n,\varepsilon}), n \geq 1\}$ is uniformly integrable. Using Fatou Lemma as well as the non-increasing property of $\widetilde{U}$, we obtain that

$$\begin{aligned}
\inf_{Z_T \in \mathcal{M}_e} E\left[\widetilde{U}(yZ_T)\right] \leq \limsup_{n\to\infty} E\left[\widetilde{U}(y\bar{Z}^{n,\varepsilon})\right] &\leq E\left[\widetilde{U}(y(1-\varepsilon)\lim_n Z^n + \varepsilon y Z_T^0)\right] \\
&\leq E\left[\widetilde{U}(y(1-\varepsilon)Y_T + \varepsilon y Z_T^0)\right] \\
&\leq E\left[\widetilde{U}(y(1-\varepsilon)Y_T)\right].
\end{aligned}$$

Repeating the previous argument and using the characterization (iv) of $AE(U) < 1$ in Lemma 3.4.10 leads to the uniformly integrability of the family $\{\widetilde{U}^{+}(y(1-\varepsilon)Y_T), \varepsilon \in (0,1)\}$. By letting ε tend to zero in the previous inequality, it follows that

$$\inf_{Z_T \in \mathcal{M}_e} E\left[\widetilde{U}(yZ_T)\right] \leq E\left[\widetilde{U}(yY_T)\right],$$

and this inequality remains valid for all $Y_T \in \mathcal{D}$. This means that

$$\inf_{Z_T \in \mathcal{M}_e} E\left[\widetilde{U}(yZ_T)\right] \leq \inf_{Y_T \in \mathcal{D}} E\left[\widetilde{U}(yY_T)\right] = \tilde{v}(y).$$

Conversely, the fact that $\mathcal{M}_e \subset \mathcal{D}$ implies that

$$\tilde{v}(y) = \inf_{Y_T \in \mathcal{D}} E\left[\widetilde{U}(yY_T)\right] \leq \inf_{Z_T \in \mathcal{M}_e} E\left[\widetilde{U}(yZ_T)\right],$$

which prove that

$$\inf_{Y_T \in \mathcal{D}} E\left[\widetilde{U}(yY_T)\right] = \inf_{Z_T \in \mathcal{M}_e} E\left[\widetilde{U}(yZ_T)\right].$$

■

Below, we will make a crucial remark which state the link between the existence of the saddle point in the method of the Lagrange multiplier and the proof of the existence and the uniqueness of the solution $\hat{X}_T^x \in \mathcal{C}(x)$ to the primal problem $v(x)$ and $\hat{Y}_T^y \in \mathcal{D}$ to the dual problem $\tilde{v}(y)$ in the Theorem 3.4.1 via the minmax theorem.

Remark 3.4.13 *From the infinite-dimensional versions of the minimax theorem B.1.3 available in the literature [ET76] are along the following lines:*

Let E, F be a couple of locally convex vector spaces in separating duality, $C \subset E$, $D \subset F$ a pair of convex subsets, and let a function $L(x,y)$ defined on $C \times D$, concave in the first and convex in the second variable, satisfying some (semi-)continuity property compatible with the topologies of E and F (which in turn should be compatible with the duality between E and F). If (at least) one of the sets C and D is compact and the other is complete, then one may show the existence of a saddle point $(\hat{\xi}, \hat{\eta}) \in C \times D$ with the property that

$$L(\hat{\xi}, \hat{\eta}) = \sup_{\xi \in C} \inf_{\eta \in D} L(\xi, \eta) = \inf_{\eta \in D} \sup_{\xi \in C} L(\xi, \eta)$$

As in [Sch00], we try to apply this theorem to the analogue of the Lagrangian implicitly encountered in the proof of Theorem 3.4.1 in 3.72. By fixing $x > 0$ and $y > 0$ let us formally write the Lagrangian in the infinite-dimensional setting

$$L^{x,y}(X_T, Y_T) = E[U(X_T)] - y\{E[X_T Y_T] - x\}, \qquad Y_T = \frac{dQ}{dP},$$

where $X_T \in \mathcal{C}(x)$ and $Y_T \in \mathcal{D}$. The sets $\mathcal{C}(x)$ and $\mathcal{D}$ are nice candidates for the Minimax Theorem to work out properly for a function L defined on $\mathcal{C}(x) \times \mathcal{D}$ because, both are closed, convex and bounded subsets of $L^0_+(P)$. But recall that we still need some compactness property to be able to localize the mini-maximizers (resp. maxi-minimizers) on $\mathcal{C}(x)$ (resp. $\mathcal{D}$). In general, neither $\mathcal{C}(x)$ nor $\mathcal{D}$ is compact (w.r.t. the topology of convergence in measure), i.e.,for a sequence $(f_n)_{n\geq 1}$ in $\mathcal{C}(x)$ (resp. $(g_n)_{n\geq 1}$ in $\mathcal{D}$) we cannot pass to a subsequence converging in measure. But $\mathcal{C}(x)$ and $\mathcal{D}$ have a property (see in Appendix in the Theorem A.3.2) which is close to compactness and in many applications turns out to serve just as well. Thus, we have

$$L^{x,y}(\hat{X}^x_T, \hat{Y}^y_T) = \sup_{X_T \in \mathcal{C}(x)} \inf_{Y_T \in \mathcal{D}} L^{x,y}(X_T, Y_T) = \inf_{Y_T \in \mathcal{D}} \sup_{X_T \in \mathcal{C}(x)} L^{x,y}(X_T, Y_T),$$

where $\hat{X}^x_T \in \mathcal{C}(x)$ is solution to the primal problem $v(x)$ and $\hat{Y}^y_T \in \mathcal{D}$ is the solution to the dual problem $\tilde{v}(y)$. Clearly $(\hat{X}^x_T, \hat{Y}^y_T)$ is the saddle point to the Lagrangian defined above, where x and y satisfy $x = I(y)$.

Chapter 4

Optimisation within Specific Markets

The portfolio expected utility maximization problem

$$\text{maximize} \quad E[U(x + \int_0^T \alpha_t \, d\, S_t)] \quad \text{over all} \quad \alpha \in \mathcal{A}(S) \tag{4.1}$$

where U is a utility function, $x > 0$ an initial capital and $\mathcal{A}(S)$ is some family of admissible self-financing strategies, was first examined by Robert Merton ([Mer71]) in a classical Black-Scholes financial market model. Exploiting the Markov structure of the model, he derived the Hamilton-Jacobi-Bellman (HJB) equation for the optimal solution more commonly known as the *value function* of the problem (4.1) and obtained a closed-form solution of this equation in the cases of exponential, logarithmic, and power utility functions. We refer the interested reader for his classical work ([Mer92]).

Thanks to the characterization of the no-arbitrage condition in terms of the existence of equivalent martingale measures obtained by Harrison et al. ([HP81]), a martingale approach using convex duality has been developed by Pliska ([Pli86]), Karatzas et al. ([KLS87]), and Cox and Huang ([CH89]). This methodology proved to be very successfully in solving portfolio optimization problems in diverse frameworks. The "Markov condition" is no longer needed. On the one hand, for general arbitrage-free and complete financial market models, it was shown that the optimal terminal wealth of the expected utility maximization problem is represented as inverse of "marginal utility" (the derivative of the utility function) evaluated at the density of the unique equivalent martingale measure. On the other hand, for incomplete markets described by the Itô processes and Brownian filtration, as was shown by He and Pearson ([HP91a, HP91b]) and Karatzas et al. ([KLSX91]), this method gives a duality characterization of optimal portfolios provided by the set of martingale measures. Their idea was to solve the dual problem of finding the suitable optimal martingale measure and permits to transform the initial dynamic problem into a static one and then to express the solution of the primal problem by using the convex duality. A definitive treatment under general semimartingale models is offered by Kramkov and Schachermayer ([KS99]) and Cvitanic, Schachermayer, and Wang ([CSW01]).

This approach principally gives a reduction of the original primal problem to the solution of the dual problem, however as shown in Section 4.2 below the explicit solution of the dual problem for general models of incomplete markets is itself a challenging task.

4.1 Examples in Complete Markets

In this section some applications of Kramkov-Schachermayer Theorem 3.4.1 are indicated in an arbitrage free and complete financial market. Mathematically speaking, the set of equivalent martingale probability measures is nonempty and reduces to a singleton:

$$\mathcal{M}_e(S) = \{P^0\}, \tag{4.2}$$

Let Z^0 denotes the martingale density process associated to the unique EMM P^0. In this context, the dual problem (3.42) is degenerate:

$$\tilde{v}(y) = E[\widetilde{U}(yZ_T^0)], \quad y > 0, \tag{4.3}$$

and the solution $\hat{Y}_T^y$ to the dual problem is clearly Z_T^0. The solution to the primal optimisation problem $v(x)$ is

$$\hat{X}_T^x = I(\hat{y}Z_T^0),$$

where $\hat{y} > 0$ is the solution to

$$E[Z_T^0 I(\hat{y}Z_T^0)] = E^{P^0}[I(\hat{y}Z_T^0)] = x.$$

In other words, the optimal terminal wealth of the expected utility maximization problem (3.7) is represented as the inverse of "marginal utility" evaluated at the density of the unique equivalent martingale measure. The wealth process $\hat{X}^x$ and the optimal portfolio $\hat{\alpha}$ are determined by

$$\hat{X}_t^x = x + \int_0^t \hat{\alpha}_s dS_s = E^{P^0}[I(\hat{y}Z_T^0)|\mathcal{F}_t], \quad 0 \le t \le T,$$

Example 4.1.1 *[Merton Model] We consider here the classical example of the Black-Scholes-Merton market model consisting of a bank account S^0 taken here as a constant $S^0 \equiv 1$ and one stock described by a geometric Brownian motion:*

$$dS_t = S_t[\mu dt + \sigma dW_t], \tag{4.4}$$

where W is a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ the natural filtration of W, $\mathcal{F}_0$ is trivial and $\mu \in \mathbb{R}$, $\sigma > 0$ are constants. We know that Black-Scholes-Merton market model is arbitrage-free and complete. Moreover, the unique equivalent martingale measure P^0 is given by its Radon-Nikodým derivative $dP^0/dP = Z_T^0$ where

$$Z_T^0 = \exp\big(-\lambda W_T - \frac{1}{2}|\lambda|^2 T\big), \quad \textit{where} \quad \lambda = \frac{\mu}{\sigma},$$

and the dynamics of the stock S under P^0 is given by

$$dS_t = S_t \sigma dW_t^0$$

where $W^0 = (W_t^0 = W_t + \lambda t)_{0 \le t \le T}$, is a P^0 standard Brownian motion. We consider the example of a constant relative risk aversion (CRRA) (also known as power utility) function:

$$U(x) = \frac{x^p}{p}, \quad 0 < p < 1, \quad \textit{for which} \quad I(y) = y^{-r}, \quad r = \frac{1}{1-p}.$$

The optimal wealth process for $v(x)$ is easily calculated:

$$\begin{aligned} \hat{X}_t^x &= E^{P^0}[(\hat{y} Z_T^0)^{-r} | \mathcal{F}_t] = \hat{y}^{-r} E^{P^0}[\exp\big(\lambda r W_T + \frac{1}{2}|\lambda|^2 r T\big) | \mathcal{F}_t] \\ &= \hat{y}^{-r} \exp\big(\frac{1}{2}(|\lambda r|^2 - |\lambda|^2 r) T\big) \exp\big(\lambda r W_t^0 - \frac{1}{2}|\lambda r|^2 t\big). \end{aligned}$$

Since $\hat{y}$ is determined by the equation $\hat{X}_0^x = x$, we obtain

$$\hat{X}_t^x = x \exp\big(\lambda r W_t^0 - \frac{1}{2}|\lambda r|^2 t\big), \quad 0 \le t \le T.$$

In order to determine the optimal control $\hat{\alpha}$, we apply Itô's formula to $\hat{X}^x$:

$$d\hat{X}_t^x = \hat{X}_t^x \lambda r dW_t^0,$$

and we compare this with

$$d\hat{X}_t^x = \hat{\alpha}_t dS_t = \hat{\alpha}_t \sigma S_t dW_t^0.$$

This provides the optimal portfolio weight $\hat{\pi}_t$, i.e. the proportion of wealth invested in risky asset S:

$$\hat{\pi}_t := \frac{\hat{\alpha}_t S_t}{\hat{X}_t^x} = \frac{\lambda r}{\sigma} = \frac{\mu}{\sigma^2(1-p)}\,.$$

The rest is being held in the risk-free asset. It is worth noting it is constant. The computation of the value function $v(x) = E[U(\hat{X}_T^x)]$ is easy, and we have explicitly

$$v(x) = \frac{x^p}{p}\,\exp\big(\frac{1}{2}\frac{\mu^2}{\sigma^2}\frac{p}{1-p}T\big)\,.$$

Example 4.1.2 (Quadratic Utility) *Let the utility function be*

$$U(x) = \mu\, x - \frac{1}{2}x^2\,,$$

we find that

$$I(y) = \mu - y.$$

As in the above example (4.1.1), the optimal wealth process for $v(x)$ is readily obtained:

$$\begin{aligned}\hat{X}_t^x &= E^{P^0}[\mu - \hat{y} Z_T^0|\mathcal{F}_t] = \mu - \hat{y}\, E^{P^0}[\exp\big(-\lambda W_T - \frac{1}{2}|\lambda|^2 T\big)|\mathcal{F}_t] \\ &= \mu - \hat{y}\,\exp\big(|\lambda|^2\, T\big)\exp\big(-\lambda W_t^0 - \frac{1}{2}|\lambda|^2 t\big).\end{aligned}$$

Since $\hat{y}$ is determined by the equation $\hat{X}_0^x = x$, we obtain

$$\hat{X}_t^x = \mu - (\mu - x)\,\exp\big(-\lambda W_t^0 - \frac{1}{2}|\lambda|^2 t\big), \quad 0 \le t \le T.$$

4.2 Examples in incomplete markets

In the setting of arbitrage-free but incomplete financial markets models, i.e. the set $\mathcal{M}_e(S)$ of equivalent martingale measures does not reduce to a singleton, and since a convex combination of two equivalent martingale measures is also an equivalent martingale measure, the set $\mathcal{M}_e(S)$ is in fact of infinite cardinality. It is very difficult if not impossible to find the explicit solution to the dual problem (3.42). However, some computations may be carried more or less explicitly in some particular models. Let us consider here the financial market model with Itô price processes and Brownian filtration as presented in Section (2.7). Observe that since we assumed the risk-premium process to be bounded, it is arbitrage-free and in general incomplete. In the notation of Section (2.7) we consider the set including $\mathcal{M}_e(S)$

$$\mathcal{M}_{loc} = \{Z_T^{\nu} \, : \, \nu \in K(\sigma)\} \supset \mathcal{M}_e(S) = \{Z_T^{\nu} \, : \, \nu \in K_m(\sigma)\}.$$

For any $\nu \in K(\sigma)$ and any wealth process $X^x = x + \int \alpha dS$, $\alpha \in \mathcal{A}(\mathcal{S})$, as an application of Itô formula, we can show that the process $Z^{\nu}X^x$ is a P-local martingale. Moreover, notice that for all $\nu \in K(\sigma)$, the bounded process $\nu^n = \nu \mathbf{1}_{|\nu| \le n}$ belongs to $K_m(\sigma)$ (remember Remark 2.7.4 where we have seen that any bounded process is in $K_m(\sigma)$), and $Z_T^{\nu^n}$ converges a.s. to Z_T^{ν}. Therefore, $M_e(S) \subset \mathcal{M}_{loc} \subset \mathcal{D}$ and from part (4) of Kramkov-Schachermayer Theorem (3.4.1), we conclude that

$$\tilde{v}(y) = \inf_{\nu \in K(\sigma)} E[\widetilde{U}(yZ_T^{\nu})], \quad y > 0. \tag{4.5}$$

The motivation and advantage of introducing the set $\mathcal{M}_{loc}$ is that it is explicit (contrary to the set $\mathcal{D}$), completely parametrized by the set of controls $\nu \in K(\sigma)$, and does not involve any assumptions about the martingale integrability as in the case of $K_m(\sigma)$, so that the stochastic control methods may be used to find a solution $\tilde{\nu}^y$ in $K(\sigma)$ to $\tilde{v}(y)$ in (4.5). In fact, if we make the additional assumption that the function

$$\xi \in \mathbb{R} \mapsto \widetilde{U}(e^{\xi}) \quad \text{is convex,}$$

which holds true if for example $x \in (0, \infty) \mapsto xU'(x)$ is increasing (both the logarithm and power utility functions satisfy the latter condition), then it is proved in Karatzas et al. [KLSX91] that for all $y > 0$, the dual problem $\tilde{v}(y)$ admits a solution $Z_T^{\hat{\nu}^y} \in \mathcal{M}_{loc}$. Moreover, we prove that for all $\nu \in K(\sigma)$ such that $E[\int_0^T |\nu_t|^2 dt] = \infty$, we have $E[\widetilde{U}(yZ_T^{\nu})] = \infty$, thus in the dual problem (4.5), we can restrict ourselves to taking the infimum over $K_2(\sigma) = \{\nu \in K(\sigma) \, : \, E[\int_0^T |\nu_t|^2 dt] < \infty\}$, and furthermore $\hat{\nu}^y \in K_2(\sigma)$. Note that this solution $Z_T^{\hat{\nu}^{\hat{y}}}$ does not belong (in general) to the set $\mathcal{M}_e(S)$. From the Kramkov-Schachermayer Theorem (3.4.1), the solution to the dual problem is then given by

$$\hat{X}_T^x = I(\hat{y} Z_T^{\hat{\nu}^{\hat{y}}})$$

where $\hat{y} > 0$ is the solution to $\operatorname{argmin}_{y>0}[\tilde{v}(y) + xy]$ and satisfying

$$E[Z_T^{\hat{\nu}^{\hat{y}}} I(\hat{y} Z_T^{\hat{\nu}^{\hat{y}}})] = x.$$

Recall that the function $I(\cdot)$ (the continuous, strictly decreasing inverse of the marginal utility function U' on $(0, \infty)$) maps $(0, \infty)$ onto itself, hence the optimal wealth process

$\hat{X}^x$ is nonnegative. Moreover we notice that the process $Z^{\hat{\nu}\hat{y}}\hat{X}^x$ is a nonnegative P-local martingale, hence a supermartingale with the property that $E[Z_T^{\hat{\nu}\hat{y}}\hat{X}_T^x] = x$. Therefore, it is a true martingale, and we finally have determined the optimal wealth process $\hat{X}^x$ as

$$\hat{X}_t^x = E\big[\frac{Z_T^{\hat{\nu}\hat{y}}}{Z_t^{\hat{\nu}\hat{y}}} I(\hat{y}Z_T^{\hat{\nu}\hat{y}})\big|\mathcal{F}_t\big], \quad 0 \le t \le T\,.$$

We apply now the results described above to two examples of utility functions.

Example 4.2.1 (Logarithmic utility function) *In this example we choose for our optimization problem (3.7) the logarithmic utility function $U(x) = \ln(x)$, $x > 0$ which plays a special role in portfolio choice, for which we have*

$$I(y) = \frac{1}{y} \quad \textit{and} \quad \widetilde{U}(y) = -\ln(y) - 1, \quad y > 0.$$

For all $\nu \in K_2(\sigma)$, we have

$$E[\widetilde{U}(yZ_T^{\nu})] = -\ln(y) - 1 + \frac{1}{2}E\big[\int_0^T (|\lambda_s|^2 + |\nu_s|^2)ds\big], \quad y > 0.$$

Therefore, the solution to the dual problem (4.5) is reached for $\nu = 0$ (independently of y) and correspond to Z_T^0. Moreover, we have

$$E[Z_T^0 I(\hat{y}Z_T^0)] = E[Z_T^0\frac{1}{\hat{y}Z_T^0}] = x\,,$$

thus the Lagrange multiplier is $\hat{y} = \frac{1}{x}$. The optimal wealth process $\hat{X}^x$ for $v(x)$, also known as the "growth optimal portfolio", is described explicitly by

$$\hat{X}_t^x = E\big[\frac{Z_T^0}{Z_t^0}\frac{1}{\hat{y}Z_T^0}\big|\mathcal{F}_t\big] = \frac{x}{Z_t^0}\,, \quad 0 \le t \le T.$$

In order to obtain the optimal control $\hat{\alpha}$ we apply Itô lemma to the above equation, and we make the identification with the dynamics $d\hat{X}_t^x = \hat{\alpha}_t dS_t$. In a financial market model written in the "Doléans-Dade exponential" form

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t),$$

we find the optimal portfolio weight $\hat{\pi}_t$, i.e. the optimal proportion of wealth invested in S:

$$\hat{\pi}_t := \frac{\hat{\alpha}_t S_t}{\hat{X}_t^x} = \frac{\mu_t}{\sigma_t^2}\,.$$

This solution is called a myopic solution, i.e. it depends only on the local behavior of the price process, in accordance with terminology adopted by Merton.

Example 4.2.2 (Power utility function) *We consider again the case of the power utility function* $U(x) = \dfrac{x^p}{p}$, $x > 0$ *with* $p \in (0,1)$ *which exhibits constant relative risk aversion (CRRA), for which we have*

$$I(y) = y^{\frac{1}{p-1}} \quad \text{and} \quad \widetilde{U}(y) = \frac{y^{-q}}{q}, \quad y > 0\,, \quad q = \frac{p}{1-p}.$$

For every $\nu \in K(\sigma)$*, we have*

$$E[\widetilde{U}(yZ_T^{\nu})] = \frac{y^{-q}}{q}\, E[(Z_T^{\nu})^{-q}], \quad y > 0.$$

From the above equation, we notice that the solution to the dual problem $\tilde{v}(y)$ *is independent of* y *and is a solution to the following problem*

$$\inf_{\nu \in K(\sigma)} E[(Z_T^{\nu})^{-q}]\,. \tag{4.6}$$

In a Markovian setting, such as a stochastic volatility model, an explicit solution to the above stochastic control problem can be derived by using the traditional dynamic programming approach. In a more general setting of Itô processes, some methods of Backward SDEs may be used (see [MT03] and the reference given there). If $\hat{\nu}$ *denote the solution to equation (4.6), the optimal wealth process is given by*

$$\hat{X}_t^x = \frac{x}{E[(Z_T^{\hat{\nu}})^{-q}]}\, E\big[\frac{(Z_T^{\hat{\nu}})^{-q}}{Z_t^{\hat{\nu}}}\big|\mathcal{F}_t\big]\,, \quad 0 \le t \le T\,.$$

Appendix A

Complements of Integration

Let us denote by $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, P)$ a filtered probability space and $L^1(\Omega, \mathcal{F}, P)$ the set of integrable random variables.

A.1 Uniform Integrability

Definition A.1.1 *(Uniformly integrable random variables) Let $(f_i)_{i\in I}$ be a family of random variables in $L^1(\Omega, \mathcal{F}, P)$. We say that $(f_i)_{i\in I}$ is uniformly integrable if*

$$\lim_{x\to\infty} \sup_{i\in I} E[|f_i| \mathbf{1}_{|f_i|\geq x}] = 0. \tag{A.1}$$

We note that any family of random variables, bounded by a fixed integrable random variable (in particular any finite family of random variables in $L^1(\Omega, \mathcal{F}, P)$ is uniformly integrable.

The following result extends the dominated convergence theorem.

Theorem A.1.2 *Let $(f_n)_{n\geq 1}$ be a sequence of random variables in $L^1(\Omega, \mathcal{F}, P)$ converging a.s to a random variable f . Then f is integrable and the convergence of f_n to f holds in $L^1(\Omega, \mathcal{F}, P)$ if and only if the sequence $(f_n)_{n\geq 1}$ is uniformly integrable. When the random variables f_n are nonnegative, this is equivalent to*

$$\lim_{n\to\infty} E[f_n] = E[f] \tag{A.2}$$

The following corollary is used in the proof of Theorem 3.2.3.

Corollary A.1.3 *Let $(f_n)_{n\geq 1}$ be a sequence of nonnegative random variables bounded in $L^1(\Omega, \mathcal{F}, P)$, i.e. $\sup_n E[f_n] < \infty$, converging a.s to a nonnegative random variable f and* *such that* $\lim_{n\to\infty} E[f_n] = E[f] + \delta$ *with* $\delta > 0$. *Then, there exists a subsequence* $(f_{n_k})_{k\geq 1}$ *of* $(f_n)_{n\geq 1}$ *and a disjoint sequence* $(A_k)_{k\geq 1}$ *of* $(\Omega, \mathcal{F})$ *such that*

$$E[f_{n_k} \boldsymbol{1}_{A_k}] \geq \frac{\delta}{2}, \qquad \forall k \geq 1. \tag{A.3}$$

Proof. A proof may be found in [P09].

The following result, due to la Vallée-Poussin, gives a practical condition for proving the uniform integrability.

Theorem A.1.4 *(la Vallée-Poussin) Let* $(f_i)_{i\in I}$ *be a family of random variables in* $L^1(\Omega, \mathcal{F}, P)$. *The following assertions are equivalent*

(i) $(f_i)_{i\in I}$ *is uniformly integrable*

(ii) *There exists a nonnegative function* φ *defined on* $\mathbb{R}_+$, $\lim_{x\to\infty} \frac{\varphi(x)}{x} = \infty$, *such that*

$$\sup_{i\in I} E[\varphi(|f_i|)] < \infty.$$

In practice, we often use the implication $(ii) \Longrightarrow (i)$. For example, by taking $\varphi(x) = x^2$, we see that any family of random variables bounded in L^2 is uniformly integrable.

A.2 Essential Supremum of a Family of Random Variables

Definition A.2.1 *(Essential supremum)*

Let $(f_i)_{i\in I}$ *be a family of real-valued random variables. The essential supremum of this family, denoted by* $\operatorname{ess\,sup}_{i\in I} f_i$ *is a random variable* $\hat{f}$ *such that*

(a) $f_i \leq \hat{f}$ *a.s., for all* $i \in I$

(b) *If* g *is a random variable satisfying* $f_i \leq g$ *a.s., for all* $i \in I$, *then* $\hat{f} \leq g$ *a.s.*

Theorem A.2.2 *Let* $(f_i)_{i\in I}$ *be a family of real-valued random variables. Then,* $\hat{f} = \operatorname{ess\,sup}_{i\in I} f_i$ *exists and is unique. Moreover, if the family* $(f_i)_{i\in I}$ *is stable by supremum, i.e. for all* i, j *in* I, *there exists* k *in* I *such that* $f_i \vee f_j = f_k$, *then there exists an increasing sequence* $(f_{i_n})_{n\geq 1}$ *in* $(f_i)_{i\in I}$

$$\hat{f} = \lim_{n\to\infty} \uparrow f_{i_n} a.s..$$

We define the essential infimum of a family of real-valued random variables $(f_i)_{i\in I}$ by: $\operatorname{ess\,inf}_{i\in I} f_i = -\operatorname{ess\,sup}_{i\in I}(-f_i)$.

A.3 Some Compactness Theorems in Probability

This first compactness result is well-known, and due to Komlos, and states that for a bounded sequence $(f_n)_{n\geq 1}$ in L^1 there is a subsequence converging in Cesaro-mean almost surely, more precisely, we have

Theorem A.3.1 (Komlos) *Let $(f_n)_{n\geq 1}$ be a sequence of random variables bounded in $L^1(\Omega, \mathcal{F}, P)$. Then, there exists a subsequence $(f_{n_k})_{k\geq 1}$ of (f_n) and a random variable f in $L^1(\Omega, \mathcal{F}, P)$ such that*

$$\frac{1}{k}\sum_{j=1}^{k} f_{n_j} \longrightarrow f, \qquad \textit{a.s when } k \textit{ goes to infinity.}$$

The following compactness theorem in $L^0(\Omega, \mathcal{F}, P)$ is very useful for deriving existence results in optimization problems in finance. It is proved in the appendix of Delbaen and Schachermayer [DS94]

Theorem A.3.2 *Let $(f_n)_{n\geq 1}$ be a sequence of random variables in $L^0(\Omega, \mathcal{F}, P)$. Then, there exists a sequence $g_n \in conv(f_n, f_{n+1}, \cdots)$, i.e. $g_n = \sum_{k=n}^{N_n} \lambda_k f_k \quad \lambda_k \in [0,1]$ and $\sum_{k=n}^{N_n} \lambda_k = 1$, such that the sequence $(g_n)_{n\geq 1}$ converges a.s. to a random variable g with values in $[0, \infty]$.*

Appendix B

Convex Analysis

Standard references for convex analysis are the books by Rockafellar [Roc70] and Ekeland and Temam [ET76]. We define $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

B.1 Semicontinuous, Convex Functions

Given a function f from open set $\mathcal{O}$ of $\mathbb{R}^d$ into $\bar{\mathbb{R}}$, we define the functions f_* and $f^* : \mathcal{O} \to \bar{\mathbb{R}}$ by

$$
\begin{aligned}
f_*(x) &= \liminf_{y \to x} f(y) := \liminf_{\varepsilon \to 0} \{f(y) : y \in \mathcal{O}, \ \ |y - x| \leq \varepsilon\} \\
f^*(x) &= \limsup_{y \to x} f(y) := \limsup_{\varepsilon \to 0} \{f(y) : y \in \mathcal{O}, \ \ |y - x| \leq \varepsilon\}.
\end{aligned}
$$

Definition B.1.1 (Semicontinuity) *Let f be a function from open set $\mathcal{O}$ of $\mathbb{R}^d$ into $\bar{\mathbb{R}}$. We say that f is lower-semicontinuous (l.s.c.) if one of the following equivalent conditions is satisfied:*

(i) $\forall x \in \mathcal{O}, f(x) \leq \liminf_{n\to\infty} f(x_n)$ *for all sequence $(x_n)_n$ converging to x.*

(ii) $\forall x \in \mathcal{O}, f(x) = f_*(x)$.

(iii) $\{x \in \mathcal{O}, f(x) \leq \lambda\}$ *is closed for all $\lambda \in \mathbb{R}$.*

We say that f is upper-semicontinuous (u.s.c.) if $-f$ is lower-semicontinuous.

We note that f is continuous on $\mathcal{O}$ if and only if f is lower and upper-semicontinuous. The function f_* is called a lower-semicontinuous envelope of f : it is the largest l.s.c. function below f . The function f^* is called a upper-semicontinuous envelope of f: it is the smallest u.s.c. function above f .

Theorem B.1.2 *A l.s.c. (resp. u.s.c.) function attains its minimum (resp. maximum) on any compact set.*

Given a convex subset C of E vector space, we recall that a function f from C into $\bar{\mathbb{R}}$ is convex if for all $x, y \in C, \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$. We say that f is strictly convex on C if for all $x, y \in C, x \neq y, \lambda \in [0,1], f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$. We say that f is (strictly) concave if $-f$ is (strictly) convex.

Theorem B.1.3 *(Minimax) Let $\mathcal{X}$ be a convex subset of a normed vector space E, compact for the weak topology $\sigma(E, E')$, and $\mathcal{Y}$ a convex subset of a vector space. Let $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a function satisfying:*

(1) *$x \longmapsto f(x,y)$ is continuous and concave on $\mathcal{X}$ for all $y \in \mathcal{Y}$.*

(2) *$y \longmapsto f(x,y)$ is convex on $\mathcal{Y}$ for all $x \in \mathcal{X}$.*

Then, we have

$$\sup_{x \in \mathcal{X}} \inf_{y \in \mathcal{Y}} f(x,y) = \inf_{y \in \mathcal{Y}} \sup_{x \in \mathcal{X}} f(x,y).$$

In the sequel, we shall restrict ourselves to the case $E = \mathbb{R}^d$. Given a convex function f from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$, we define its domain by

$$dom(f) = \{x \in \mathbb{R}^d : f(x) < \infty\},$$

which is a convex set of $\mathbb{R}^d$. We say that a convex function f from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ is proper if it never takes the value $-\infty$ and if $dom(f) \neq \emptyset$.

We have the following continuity result for convex functions.

Proposition B.1.4 *A proper convex function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ is continuous on the interior of its domain.*

We focus on the differentiability of convex functions.

Definition B.1.5 *(Subdifferential) Given a convex function f from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$, we define the subdifferential of f in $x \in \mathbb{R}^d$, denoted by $\partial f(x)$ as the set of points y in $\bar{\mathbb{R}}$ such that*

$$f(x) + y.(z - x) \le f(z), \qquad \forall z \in \mathbb{R}^d.$$

Proposition B.1.6 *Let f a convex function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$.*

(1) *If f is finite and continuous at $x \in \mathbb{R}^d$, then $\partial f(x) \neq \emptyset$.*

(2) *f is finite and differentiable at $x \in \mathbb{R}^d$ with gradient $Df(x)$ if and only if $\partial f(x)$ is reduced to a singleton and in this case $\partial f(x) = \{Df(x)\}$.*

B.2 Fenchel-Legendre Transform

Definition B.2.1 *(Polar functions) Given a convex function f from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$, we define the polar (or conjugate) of f as the function $\tilde{f}$ from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ where*

$$\tilde{f}(y) = \sup_{x \in \mathbb{R}^d} \{x.y - f(x)\}, \qquad y \in \mathbb{R}^d.$$

When f is convex, we also say that $\tilde{f}$ is the Fenchel-Legendre transform of f. It is clear that in the definition of $\tilde{f}$ we may restrict in the supremum to the points x lying in the domain of f. The polar function $\tilde{f}$ is defined as the pointwise supremum of the affine functions $y \to x.y - f(x)$. Thus, it is a convex function on $\mathbb{R}^d$.

We may also define the polar function of a polar function. We have the following bipolarity result.

Theorem B.2.2 *(Fenchel-Moreau) Let f be proper, convex l.s.c. function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ and $\tilde{f}$ its Fenchel-Legendre transform. Then,*

$$f(x) = \sup_{y \in \mathbb{R}^d} \{x.y - \tilde{f}(y)\}, \qquad x \in \mathbb{R}^d.$$

In other words, we have $f = \tilde{\tilde{f}}$.

We state the connection between differentiability and polar functions.

Proposition B.2.3 *Let f be proper, convex l.s.c. function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ and $\tilde{f}$ its Fenchel-Legendre transform. Then, for all $x, y \in \mathbb{R}^d$ we have the following equivalence*

$$y \in \partial f(x) \Longleftrightarrow x \in \partial \tilde{f}(y) \Longleftrightarrow f(x) = x.y - \tilde{f}(y).$$

Proposition B.2.4 *Let f be proper, convex l.s.c. function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ strictly convex on the int$(dom(f))$. Then, the Fenchel-Legendre transform $\tilde{f}$ is differentiable on int$(dom(\tilde{f}))$. Furthermore, if f is differentiable on int$(dom(f))$, then the gradient of f, Df, is one-to-one from int$(dom(f))$ into int$(dom(f))$ with $Df = (D\tilde{f})^{-1}$ and $\tilde{f}$ is strictly convex on int$(dom(\tilde{f}))$.*

Appendix C

Some Results from Stochastic Analysis

Theorem C.0.5 ***(Itô representation theorem)***

Assume that $\mathbb{F}$ is the natural (augmented) filtration of a standard d-dimensional Brownian motion $W = (W^1, \cdots, W^d)$. Let $M = (M_t)_{t\in\mathbb{T}}$ be a càd-làg local martingale. Then there exists $\alpha = (\alpha^1, \cdots, \alpha^d) \in L^2_{loc}(W)$ such that

$$M_t = M_0 + \int_0^t \alpha_u \cdot dW_u = M_0 + \sum_{i=1}^{d} \int_0^t \alpha_u^i dW_u^i, \qquad t \in \mathbb{T} \ \ a.s. \tag{C.1}$$

Proof. For a proof, the interested reader refer to [KS98], Section 3.4 in Chapter 3.

Theorem C.0.6 *Let $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ be a filtration satisfying the usual conditions, and $X = (X_t)_{t\in T}$ be a supermartingale. Then X has a càd-làg modification if and only if the mapping $t \in T \to E[X_t]$ is right-continuous (this is the case in particular if X is a martingale). Moreover, in this case, the càd-làg modification remains a supermartingale with respect to $\mathbb{F}$.*

Proof. For a proof, the reader refer to [KS98], Theorem 3.13 in Ch. 1.

Bibliography

[BI73] J.-M. Bismut, *Conjugate convex functions in optimal stochastic control*, Journal of Mathematical Analysis and Applications, **44**, 1973, 384–404. 1

[BI75] J.-M. Bismut, *Growth and optimal intertemporal allocation of risks*, Journal of Economic Theory, **10**, 1975, 239–257. 1

[BB83] K. P. S. Bhaskara Rao and M. Bhaskara Rao, *Theory of Charges*, Academic Press, London, 1983.

[BS99] W. Branath and W. Schachermayer, *A bipolar theorem for subsets of* $L^0_+(\Omega, \mathcal{F}, \mathcal{P})$, Séminaire de Probabilités **XXXIII** (1999), 349–354.

[CH89] J. C. Cox and C. F. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, J. Economic Theory **49** (1989), 33–83. i, 1, 79

[CH91] J. C. Cox and C. F. Huang, *A variational problem arising in financial economics*, J. Math. Econ **20** (1991), 465–487. i

[Chu74] K.-L. Chung, *A Course in Probability Theory*, Second Edition, Academic Press, New York-London, 1974, Probability and Mathematical Statistics, Vol. 21.

[CK92] J. Cvitanić and I. Karatzas, *Convex duality in constrained portfolio optimization*, Ann. Appl. Probab. **2** (1992), no. 4, 767–818.

[CSW01] J. Cvitanić, W. Schachermayer, and H. Wang, *Utility maximization in incomplete markets with random endowment*, Finance and Stochastics **5** (2001), 237–259. 80

[Cuo97] D. Cuoco, *Optimal consumption and equilibrium prices with portfolio constraints and stochastic income*, J. Econom. Theory **72** (1997), 33–73.

[DM82] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential, Volume B: Theory of Martingales*, North-Holland, Amsterdam, 1982.

[DS94] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen **300** (1994), 463–520. 17, 88

[DS95] F. Delbaen and W. Schachermayer, *The existence of absolutely continuous local martingale measures*, Ann. Appl. Probab. **5** (1995), no. 4, 926–945. ii, 15, 17

[DS98] F. Delbaen and W. Schachermayer, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Mathematische Annalen **312** (1998), 215–250. ii, 15, 17

[DS99] F. Delbaen and W. Schachermayer, *A compactness principle for bounded sequences of martingales with applications*, Proceedings of the Seminar of Stochastic Analysis, Random Fields and Applications, 1999. 17

[ET76] I. Ekeland, and R. Temam, *Convex Analysis and Variational Problems*, North Holland, 1976. 42, 77, 89

[FK97] H. Föllmer and D. Kramkov, *Optional decomposition under constraints*, Probability Theory and Related Fields **109** (1997), 1–25. 13

[HP81] J. M. Harrison and S. R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Process. Appl. **11(3)** (1981), 215–260. 14, 79

[HP91a] H. He and N. D. Pearson, *Consumption and portfolio policies with incomplete markets and short-sale constraints: the finite-dimensional case*, Mathematical Finance **1** (1991), 1–10. 1, 79

[HP91b] H. He and N. D. Pearson, *Consumption and portfolio policies with incomplete markets and short-sale constraints: the infinite-dimensional case*, Journal of Economic Theory **54** (1991), 259–304. 1, 79

[Jac79] J. Jacod, *Calcul Stochastique et Problèmes de Martingales (Lecture Notes in Mathematics 714)*, Springer, Berlin Heidelberg New York, 1979.

[KJP98] N. El Karoui and M. Jeanblanc-Picqué, *Optimization of consumption with labor income*, Finance and Stochastics **2** (1998), 409–440.

[KLS87] I. Karatzas, J. P. Lehoczky, and S. E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM Journal on Control and Optimization **25** (1987), 1557–1586. i, 1, 5, 79

[KLSX91] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM Journal on Control and Optimization **29(3)** (1991), 702–730. 1, 79, 83

[KQ95] N. El Karoui and M.-C. Quenez, *Dynamic programming and pricing of contingent claims in an incomplete market*, SIAM Journal on Control and Optimization **33/1** (1995), 29–66. ii, 13, 15, 18, 30

[Kra96] D. Kramkov, *Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets*, Probability Theory and Related Fields **105** (1996), 459–479. ii, 13, 18

[KR07] I. Klein, and L. C. G. Rogers, *Duality in optimal investment and consumption problems with market frictions* Math. Finance **17** (2007), no. 2, 225–247. ii, 2

[KS91] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Springer-Verlag, New York, 1991.

[KS98] I. Karatzas and S. E. Shreve, *Methods of Mathematical Finance*, Springer Verlag, New York, 1998. 47, 93

[KS99] D. Kramkov and W. Schachermayer, *A condition on the asymptotic elasticity of utility functions and optimal investment in incomplete markets*, Annals of Applied Probability **9** (1999), no. 3, 904–950. i, 48, 50, 80

[LS91] P. Lakner and E. Slud, *Optimal consumption by a bond investor: the case of random interest rate adapted to a point process*, SIAM Journal on Control and Optimization **29** (1991), no. 3, 638–655.

[MT03] M. Mania and R. Tevzadze, *Backward stochastic PDE and imperfect hedging*, Int. J. Theor. Appl. Finance **6** (2003), no. 7, 663–692. 35, 85

[Mer69] R. C. Merton, *Lifetime portfolio selection under uncertainty: the continuous-time case*, Rev. Econom. Statist. **51** (1969), 247–257. i, 1

[Mer71] R. C. Merton, *Optimum consumption and portfolio rules in a conitinuous-time model*, J. Economic Theory **3** (1971), 373–413. i, 1, 79

[Mer92] R. C. Merton, *Continuous-time Finance*, Revised Edition, Oxford, Basil Blackwell, 1992. 79

[P09] H. Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer-Verlag Berlin Heidelberg, 2009. i, ii, 11, 87

[Pli86] S. R. Pliska, *A stochastic calculus model of continuous trading: optimal portfolio*, Math. Oper. Res. **11** (1986), 371–382. i, 1, 79

[Pr07] J. L. Prigent, *Portfolio Optimization and Performance Analysis*, New York, 2007. 6

[Pro90] Ph. Protter, *Stochastic Integration and Differential Equations*, Springer Verlag, New York, 1990.

[Roc70] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970. 42, 89

[R03] L. C. G. Rogers, *Duality in constrained optimal investment and consumption problems: a synthesis*, Paris-Princeton Lectures on Mathematical Finance, 2002, 95–131, Lecture Notes in Math., **1814**, Springer, Berlin, 2003. ii, 2

[Sch86] M. Schwartz, *New proofs of a theorem of Komlós*, Acta Math. Hung. **47** (1986), 181–185.

[Sch00] W. Schachermayer, *Optimal investment in incomplete financial markets*, to appear in Proceedings of the First World Congress of Bachelier Society, Paris (2000). 78

[Shi96] A. N. Shiryaev, *Probability*, Second Edition, Springer-Verlag, New York, 1996. Translated from the first (1980) Russian edition by R. P. Boas.

[Str85] H. Strasser, *Mathematical Theory of Statistics*, De Gruyter, Berlin-New York, 1985.

[RSF08] S.T. Rachev, S. V. Stoyanov, F. J. Fabozzi, *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures*, (Frank J. Fabozzi Series), 2008. 6

[Woj96] P. Wojatszcyk, *Banach Spaces for Analysts*, Cambridge University Press, Cambridge, 1996.

[Xu90] G.-L. Xu, *A duality method for optimal consumption and investment under short-selling prohibition*, Ph.D. thesis, Department of Mathematics, Carnegie-Mellon University, 1990.

[YH52] K. Yosida and E. Hewitt, *Finitely additive measures*, Transactions of the American Mathematical Society **72** (1952), 46–66.

[Žit00] G. Žitković, *A filtered version of the bipolar theorem of Brannath and Schachermayer*, Journal of Theoretical Probability **15** (2002), 41–61.